Multiband Image Fusion via Linear Mappings

Toby Nicholas Matheson

A Thesis submitted for the degree of Doctor of Philosophy

School of Computing Sciences

University of East Anglia

September 2019

Contents

List of Figures

List of Tables

Publications

G. D. Finlayson and T. N. Matheson. A closed form solution for the brightness preserving colour to greyscale image conversion. In Colour Imaging Conference, 2012

Acknowledgements

I've had many primary supervisors over the years. I would like to thank Dr Wenjia Wang who took on the difficult task of taking over my supervision while I was entering my registration year. His guidance and tuition on how to produce a PhD thesis as well as his moral and emotional support leading up to my viva was invaluable to me and it is unlikely that I would have passed without him.

Secondly I would like to thank Dr Mark Fisher for allowing me the intellectual freedom and space for me to pursue the line of research and PhD that I wanted to do. Although I lacked gratitude at the time, I would like him to know that his kindness and support was not lost on me and that it meant a great deal.

It has to be said that all of this would not have been possible without my first supervisor Prof. Graham Finlayson. I wish to thank him for originally accepting me onto the PhD program and for introducing me to computational geometry and the brightness preserving greyscale problem.

The success of my academic journey was augmented by other members of my supervisory team and University faculty. I am eternally grateful for the support of supervisory team member Dr Michal Mackiewicz. I've lost count of the number of times that I made use of his open door policy to discuss whatever was plaguing my thoughts at the time. From linear programming to intersecting half-spaces or just general day-to-day issues, he always made time for me.

I would like to thank computing science faculty member Dr Pierre Chardaire for counselling me on mathematical notation and engaging with me on any little problem I was chewing on at the time. I still fondly recall the hours lost in his office while we looked at solving quartic equations.

I would like to thank Dr Robert Whittaker from the school of mathematics for his guidance

in set theory notation and for his assistance in arriving at the solutions (Eqs. (3.34) - (3.38)) in Chapter 3. The speed at which he grasped what I was working on was a truly humbling experience and it would be a pleasure to work with him in the future should our paths ever cross again.

The Colour Lab here at the UEA hosts some of the best computer scientists I've had the pleasure of meeting. Of these I would like to thank former member Dr Roberto Montagna, for always being available to assist me with general questions in MATLAB and imaging.

I would like to say a very special thank you to my internal and external examiners, Dr Ben Milner and Dr Gerald Schaefer. I would like to thank them for making my viva experience engaging and constructive in addition to diligently producing a set of corrections that has made this final version of my thesis many times better than what I originally submitted.

The low computational complexity of my PhD meant that I did most if not all of it on my laptop. This however did not stop me from making use of our schools technicians, namely Matthew Ladd and Peter Trollope. Whether it be simple printing issues to repairing my Apple Mac power adapter, they were always on hand and willing to assist me.

I owe a lot to my mental health wellbeing mentor and the larger student services team here at the UEA. I thank them for equipping me with the tools on how to cope with stress, anxiety and financial difficulties.

A big thank you to Dr Katharina Huber and the postgraduate office members Catherine Taylor and Sonia Fagan for always being supportive and hasty with my numerous last minute concession requests.

Lastly I want to thank my father for helping me through what I feel has been the hardest time of my life. Without his intermittent financial support, I would never have been able to complete this journey that I chose to embark upon.

Abstract

Despite the development of more sophisticated airborne systems that are equipped with hyperspectral sensors, the more cost effective commercially sponsored multispectral satellite systems are still in use today. Consequently, the research topic of multispectral panchromatic image sharpening is still active, whose sole purpose is to produce high spatial resolution imagery while preserving the spectral integrity of the original spectral image. The component substitution injection scheme is still a foundation for many of today's techniques. Working under the assumption that a low-pass filtered panchromatic image can be constructed from a linear mapping of the spectral bands, an unsharp mask can be created and fused into the original multispectral image. The challenge in deciding how the band weights are computed is typically solved using linear regression, however, this produces a greyscale with different global statistics to the standard intensity mapping thus altering the spectral properties of the pan-sharpened image. We tackle this problem with a different approach; that is to produce a method of mapping a four band image into a real greyscale that preserves the global statistics of the standard intensity mapping all the while increasing the correlation with the panchromatic. Starting with colour to greyscale mappings, we produce a closed form solution to a maximum variance greyscale subject to preserving image mean. By exploiting the cubic constraints on the band weights, we reduce the time complexity from that of quadratic programming to one which is limited by sorting three numbers. We further expand on this optimisation to produce greyscales with maximum image variance which has application to multi-banded images for dimensions limited by the ability to compute a convex-hull. Lastly we adapt our solution by introducing a quadratic constraint on image variance. This method is based on a novel solution to the geometric intersection of a hyper-ellipse with hyper-planes. Lastly we further this solution to finding the band weights that will enhance the correlation of our mapping to the panchromatic image. Post image evaluation our technique was seen to improve on classical component substitution methods.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Chapter 1

Introduction

1.1 Motivation

The broad definition of remote sensing includes many forms of data collection that have been carried out at a distance as opposed to in-situ (Schowengerdt, 2007). This encompasses vision, astronomy, space probes, most medical imaging and sonar (Schott, 2007). For our purposes, we shall solely associate this term to that of digital image acquisition by aircraft and satellites.

The year 1972 saw the release of the United States Geological Survey (USGS) & North American Space Association (NASA) Landsat Multispectral Scanner System (MSS) satellite, which possessed a sensor capable of capturing surface reflectance data from Earth over four regions (multispectral bands each with a bandwidth of about 100 nm) of the electromagnetic (EM) spectrum. Since then, we have seen the release of four additional MSS (Landsat 2 - 5, the last two of which also using the Thematic Mapper (TM) sensor), the Enhanced Thematic Mapper Plus (ETM+) sensor (Landsat 7) and most recently Landsat 8, which is equipped with the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS), capable of collecting image data for nine short-wave bands and two long-wave thermal bands respectively.

There have also been seven French "Satellite Pour l'Observation de la Terre" (SPOT) systems as well as the commercially (Digital Globe) developed IKONOS, QuickBird and OrbView earth observation platforms. The optics that collect spectral information on many of these commercial satellite platforms generally comprise around four to ten spectral band sensors i.e. Coastal

Aerosol, Blue, Green, Red, Near-Infrared, Short-Wave Infrared 1&2, Cirrus and Thermal Infrared 1&2 ($\boldsymbol{CA}, \boldsymbol{B}, \boldsymbol{G}, \boldsymbol{R}, \boldsymbol{NIR}, \boldsymbol{SWIR}, \boldsymbol{C}, \boldsymbol{TIRS}$) (Barsi et al., 2014). These image sets are referred to as *multispectral*, whereas satellites such as the Advanced Visible/Infrared Imaging Spectrometer (AVIRIS) utilise hundreds of sensors (with approximately 10 nm wide bandwidths) that collect data for 400 to 2400 nm of the electromagnetic spectrum and are referred to as *hyperspectral*. Ghassemian (2016) states that it is not unheard of to add further classification by referring to systems that capture more than ten bands (with bandwidths between 50 - 10 nm) in the EM spectrum as *superspectral* and sensors with bandwidths of less that or equal to 1 nm as *ultraspectral*.

The bandwidth of these spectral bands is confined to the window of electromagnetic radiation the sensors can detect; the narrower this bandwidth, the greater the sensors *spectral* resolution. The *spatial* resolution of a sensor depends on its Instantaneous Field Of View (IFOV): the ground area captured by the sensor at any given instant of time (Campbell, 2006). Digitally this can be thought of as the area of the ground captured by one pixel. For a given number of pixels, the finer the IFOV is, the higher the spatial resolution; the greater the clarity of the high-frequency information available in an image (in the absence of sensor noise).

The design of multispectral sensors with increasingly higher spatial resolutions is limited by technical constraints of onboard storage and bandwidth transmission from the satellite to the Earth's surface. For this reason, Earth observation platforms are equipped with a single high spatial, low spectral resolution *panchromatic* sensor that is sensitive to a large portion of the EM spectrum from the visible to the thermal infrared. Images acquired from the panchromatic sensor are often referred to as the "panchromatic image" $\boldsymbol{P}$ and it provides a single measure of reflectance of the scene over the wavelengths that the panchromatic sensor can detect. In addition to data handling constraints, remote sensing systems are subject to the trade-off between IFOV and signal-to-noise ratio (SNR). Since multispectral, and to a greater extent hyperspectral, sensors have reduced spectral bandwidths compared to panchromatic sensors, they typically have for a given IFOV a reduced spatial resolution in order to collect more photons and preserve the image SNR. With these constraints in mind, generating high spatial and spectral resolution imagery are not immediately available via image acquisition by remote sensing but instead by fusion methods. This concept is known as multispectral or multi-sensor merging, fusion or *pan-sharpening* which has been defined by Amro et al. (2011) as “a pixel-level fusion technique used to increase the spatial resolution of the multispectral image, while preserving the spectral information in the multispectral image”.

1.2 Research Problem

The *classical techniques* of pan-sharpening rely on *Component Substitution* (CS) and include the Intensity-Hue-Saturation (IHS), Principal Component Analysis (PCA) and Gram Schmidt (GS) methods (Amro et al., 2011). Pan-sharpening based on CS is one of the most popular methods to employ due to its simplicity, computational speed and the ability to inject high level of spatial detail. The technique relies upon swapping out the low-resolution spatial information (or part of it) with that of the high-resolution $\boldsymbol{P}$. How the low-resolution spatial information is acquired from the colour image depends on the transformation done prior to the substitution. The transformation can be viewed as a projection of the image from RGB vector space to a component space that separates the spatial from the spectral information. After the substitution has been made, the inverse transform provides the final pan-sharpened image.

CS is an incredibly simple method for fast image pan-sharpening. Undesirably, this simplicity comes at a cost; CS methods suffer the most from a significant problem in the field of image fusion: *spectral distortion*. Spectral distortion describes a whole range of unwanted characteristics that can be present in the fused image, from visual artefacts/errors such as contrast inversion to quantifiable deviations from a statistical norm or metric. It has been identified that one of the main causes of these distortions is caused by the spectral mismatch between the multispectral and panchromatic sensors, see Figure. 1.1. Here we can see the dispersive nature of each sensor; they are not equally sensitive across their respective bandwidths, as seen by the normalised peaks in the spectra. Among the ‘tails’ of the band sensor responses there exists redundant overlapping, with the most severe being between the blue and green bands. The overlapping between adjacent bands diminishes with increasing wavelength, to the extent that the red and near-infrared ($\boldsymbol{NIR}$) being almost entirely vacant. The panchromatic band is depicted by the black line where its low spectral resolution can be seen by its broad bandwidth. The spectral mismatch can be identified by the panchromatic sensor being overly sensitive in the infrared and

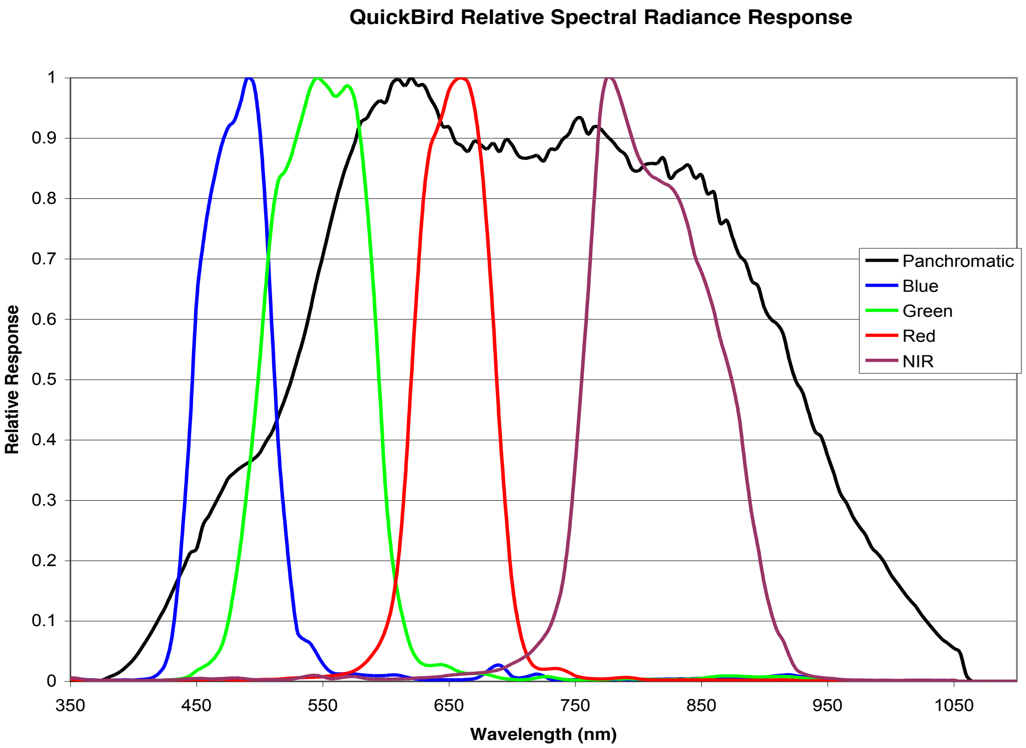


Figure 1.1: QuickBird Relative Spectral Response (Dig, 2019).

not sensitive enough in the blue; as depicted by the sensors response to EM-radiation post 950 nm and its rapid decline at wavelengths smaller than 450 nm. A good example as to how this mismatch of the spectral responses can affect a fused image can be made when considering the spectral characteristics of vegetation (Zhang, 2004); vegetation appears of relatively high reflectance in the near-infrared and panchromatic bands, while its reflectance is low in the visible. The implications of this would be seen post CS, where the brightness levels in the pan-sharpened image with areas possessing vegetation appearing brighter than the original multispectral image. In addition to introducing this offset in brightness, using the panchromatic image in a CS can lower the correlation between the bands (Thomas et al., 2008) which can lead to contrast inversion.

1.2.1 Image Pre-processing

To mitigate the introduction of spectral distortions in the pan-sharpened image, it is usual for the panchromatic image to undergo standard (as found in many text-books on the subject such as Gonzalez and Woods, 2008) image pre-processing prior to fusion. One of these pre-processing steps aims to preserve the spectral characteristics of the multispectral image, by matching the statistical mean and standard deviation of the panchromatic to that of the component it is replacing (Amro et al., 2011). For IHS pan-sharpening this would be the intensity image $\boldsymbol{I}$: a single band image computed by averaging across the multispectral bands. In the pan-sharpening literature this process is known as *histogram matching* which is a misnomer, it is more accurate to refer to it as *statistical matching.* The remaining two pre-processing stages include *image registration*; the process of making the pixels of two or more scenes of the same geographical region precisely coincide with the same points on the ground (Zitová and Flusser, 2003) and *resampling* (upsampling), a means of changing (improving) the image resolution using interpolation techniques.

Although $\boldsymbol{P}$ is usually statistically matched prior to CS, significant spectral distortion still remain in most pan-sharpened images. Consider IHS pan-sharpening, where the global statistics of $\boldsymbol{P}$ are matched to those of $\boldsymbol{I}$. Globally the statistics have been matched, however at the local level level we could have introduced or exasperated deviations with $\boldsymbol{I}$. As Zhang (2010) has described it: "To develop an image fusion method preserving the spectral characteristics, the high resolution image has to sharpen the multispectral image without adding new grey level information to its spectral components; the local means must ideally be preserved between the panchromatic and the intensity image". Additionally, there could still exist areas where there is a lack of correlation between the $\boldsymbol{I}$ and $\boldsymbol{P}$ components which leads to areas of contrast inversion. This brings to light a fundamental assumption behind the CS technique; there exists a linear relationship between the multispectral and panchromatic bands. If this assumption were true then it would be possible to construct the panchromatic image from a linear combination of the multispectral bands.

1.2.2 Relative Spectral Contribution

A single band image represents the sensor response captured at a pixel which when viewed will appear composed of shades of grey with black representing the lowest intensity and white the highest. For this reason they are usually referred to as *greyscale* images. This idea of constructing a greyscale panchromatic-like image (known either as a synthetic/approximate panchromatic image) using a linear combination of the multispectral bands forms a new group of pan-sharpening techniques known as *Relative Spectral Contribution* (RSC). Although RSC techniques do not directly pan-sharpen multispectral images, they are usually affiliated with methods where a linear combination of the bands is used to provide the spatial detail over that of $\boldsymbol{I}$. Constructing a synthetic panchromatic image $\boldsymbol{J}$ using a linear combination of the bands can be formulated as

$$\boldsymbol{J} = \sum_{1}^{d} w_n \boldsymbol{M}_n, \tag{1.1}$$

where d is the total number of bands being used in the fusion, w_n is the n^{th} band weight and $\boldsymbol{M}_n$ is the n^{th} multispectral band. For example, the classic IHS transform uses Eq. (1.1) to compute the intensity mapping $\boldsymbol{J} = \boldsymbol{I}$ from the $\boldsymbol{R}$, $\boldsymbol{G}$ and $\boldsymbol{B}$ bands by setting $w = \frac{1}{3}$. This leads to the generalisation of intensity where any number of multispectral bands by setting $w = \frac{1}{d}$.

To illustrate the flaws of equally weighting the bands we must look back to the sensor response. Ideally, the multispectral bands $(\boldsymbol{B}, \boldsymbol{G}, \boldsymbol{R}, \boldsymbol{NIR})$ from Figure. 1.1 should be disjoint, equally sensitive and together form a continuous response that covers the range of the panchromatic sensor. This idealised conceptual sensor response is shown in Figure. 1.2. Under such circumstances the intensity mapping would produce an image that possesses the exact same spectral properties of that collected from the panchromatic sensor. From Figure. 1.1 this is clearly not the case; there is overlap between the $\boldsymbol{B}$ and $\boldsymbol{G}$ bands in addition to spectral mismatch of the $\boldsymbol{B}$ and $\boldsymbol{NIR}$ bands with $\boldsymbol{P}$. This illustrates the problem of using an equally weighted average and how fusion schemes that don't account for these spectral mismatches can cause distortions in the pan-sharpened image. Nowadays the occurrence of these distortions are reduced by computing $\boldsymbol{J}$ in such a way that the spectral response of the sensor is considered, by weighting the contributions from the multispectral bands appropriately. This is the general premise behind the RSC approach. This has been discussed by Aiazzi et al. (2007) where they

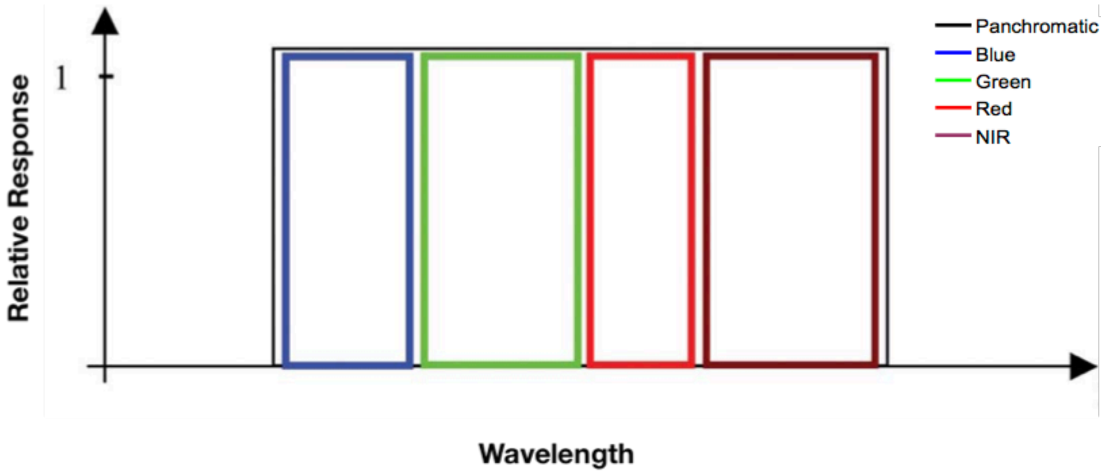


Figure 1.2: Idealised QuickBird spectral response of the multispectral bands (Thomas et al., 2008).

developed a method of determining the weights of Eq. (1.1) by minimising the square error between their $\boldsymbol{J}$ and $\boldsymbol{P}$ using linear regression.

Currently there is no linear relationship between the multispectral and panchromatic sensors. This means that the problem with computing $\boldsymbol{J}$ that adopts the approach of maximising the correlation with $\boldsymbol{P}$ is that it does not fully address the problems that using $\boldsymbol{I}$ can cause. Ultimately what we we desire is the intensity image but at a higher resolution (the ARSIS concept: Ranchin et al., 1996) and it is this objective that pan-sharpening techniques now seek to satisfy, while having as low a time-complexity as possible.

To attempt to achieve this goal, further mitigation of the spectral distortions with global CS/RSC techniques have been addressed in frequency domain methods. It is well known that the spatial information of the panchromatic image is mostly carried by its high-frequency components, while the spectral information is mostly carried by its low-frequency components (Fonseca et al., 2011, Guo et al., 2010). Performing a global CS can dramatically alter the low-frequency components of the original multispectral image, which can produce distortions in the spectral bands (Zhou et al., 1998). The basic solution to this would be to extract the high-frequency components from $\boldsymbol{P}$ and inject them into the $\boldsymbol{M}$ bands, a form of pan-sharpening known as *High-Pass Frequency* (HPF) injection (Chavez et al., 1991). With regards to the RSC method, Thomas et al. (2008) states that “a filtering operation applied on the panchromatic image is implicitly required" when using the $\boldsymbol{P}$ in a linear combination of multispectral bands.

1.3 Aims and Scope

The scope of this thesis is limited to constrained linear mappings and their application to panchromatic image sharpening. With respect to the intensity image $\boldsymbol{I}$, our final aim is to produce a method of linearly mapping four spectral bands into a single band greyscale such that the global statistics are preserved all the while having a greater correlation with the panchromatic. To accomplish this we complete the following objectives:

1. We produce a closed form solution to linearly mapping a three band (colour) image into a normalised single band (greyscale) such that the statistical mean of $\boldsymbol{I}$ is preserved.
2. Using psychophysics, we evaluate our greyscales against those from competing techniques.
3. By applying a constraint on variance, we further develop our method for use in panchromatic image sharpening. To this end we develop a novel solution to the intersection of a hyper-plane with a hyper-ellipsoid.
4. Using linear regression, we produce a novel method of producing weights that increase the correlation of $\boldsymbol{J}$ subject to our constraints on the global statistics.
5. We evaluate our pan-sharpened imagery against other methods found in the literature.

1.4 Thesis Outline

Computing a greyscale $\boldsymbol{J}$ by linearly mapping multiband data is a form of dimensionality reduction. Consequently this is a natural place for Chapter 2 to begin the narrative for this thesis: a literature review in colour-to-greyscale mappings.

Further to this review, Chapter 3 shows how we expand on and develop a closed form solution to a published brightness preserving colour-to-greyscale method that previously relied on a quadratic programming algorithm to solve.

Chapter 4 finishes our work in colour-to-grey by presenting our psychophysical evaluation experiment. Here we gauge the overall preference of our colour-to-greyscale conversion against competing techniques.

Chapter 5 sees us moving onto the research field of panchromatic image sharpening with a literature review.

In Chapter 6 we demonstrate the strength of the model by introducing a novel constraints on image variance. Additionally we show how we can produce a mapping that has a greater correlation with the panchromatic image while satisfying our constraints.

Chapter 7 is our panchromatic sharpening evaluation chapter and shows how using our greyscale we can improve on the classical methods.

Lastly we finish with a conclusion and further work discussion in Chapter 8.

Chapter 2

Colour-to-Greyscale: Literature Review

The research problem in panchromatic image sharpening is how one can reproduce the edge information of the panchromatic in the multispectral image. This problem bears strong similarity to that found in the field of colour-to-greyscale conversion, where the research community attempts to reproduce the edge information of the colour image in their greyscale mappings. Researching colour-to-greyscale methodologies is therefore the natural place to begin our research to the pan-sharpening problem.

In this chapter we shall present our literature review for the research topic of colour-to-greyscale. We shall briefly introduce the research problem before dividing the review into the four sections. The first two sections 1) luminance-chromaticity and 2) RGB colour cube based mappings deal with the current state of the art. The third section we shall look deeper and analyse their compatibility to panchromatic image sharpening. Lastly we shall finish with a conclusion.

2.1 Introduction

Linearly mapping a multiband image down to a single band is a dimensionality reduction problem. Dimensionality reduction in imaging is well documented in the field of colour-to-greyscale, where a three band image is reduced to one. The challenge concerns preserving the information content of the data as one reduces its dimensionality. This problem is generally solved by trying to find a greyscale that possesses the same contrast as that of the original colour image.

Contrast in an image is the difference in colour or brightness between two points that make them distinguishable from each other. Due to the discrete nature of digital images, the minimum incremental spatial distance between two potentially different loci of information is 1, i.e. neighbouring pixels i and j. This allows local contrast to be expressed as an approximation of a gradient, which for our greyscale can be written as $\nabla \boldsymbol{J}_{i,j}$. Replicating the contrast of the colour image in the greyscale can be posed in the image gradient domain using the cost function (Bhat et al., 2008, Jin and Süsstrunk, 2017)

$$\min \sum_{(i,j)\in\kappa} \left\| \nabla \boldsymbol{J}_{i,j} - \nabla \boldsymbol{M}_{i,j} \right\|^2 , \tag{2.1}$$

where κ represents the set of all pixel pairs, $\nabla \boldsymbol{J}_{i,j}$ contain the gradients for the single band greyscale

$$\nabla \boldsymbol{J}_{i,j} = \boldsymbol{J}_i - \boldsymbol{J}_j \tag{2.2}$$

and $\nabla \boldsymbol{M}_{i,j}$ the gradients for the three band colour image. Since gradients are scalar quantities and pixel information for multiband imagery are represented by vectors, the resultant Euclidean vector magnitude across the three bands are used to define the differences between pixels. For example, if we have a pixel at location i and another at location j then an element in the matrix $\nabla \boldsymbol{M}$ would be

$$\nabla \boldsymbol{M}_{i,j} = \oplus_{i,j} \sqrt{(\boldsymbol{R}_i - \boldsymbol{R}_j)^2 + (\boldsymbol{G}_i - \boldsymbol{G}_j)^2 + (\boldsymbol{B}_i - \boldsymbol{B}_j)^2}, \tag{2.3}$$

where $\boldsymbol{R}_i, \boldsymbol{G}_i, \boldsymbol{B}_i$ represents a vector components from the red, green and blue bands. As all Euclidean magnitudes are positive, a sign function $\oplus$ is needed to define the gradient direction (positive or negative), which visually means that the brightness ordering between the pixels is an accurate representation of what would be seen in the colour image. Typically this function produces signs based on the gradient directions from particular single band image i.e. the luminance image from either the RGB colour cube model $\boldsymbol{L}$ or L^* from the perceptually uniform CIE $L^*a^*b^*$ colour space.

Minimising the objective function described by Eq. (2.1) is a known problem in multivariate calculus, where the desired gradients $\nabla \boldsymbol{J}$ is as close as possible to the given gradient field $\nabla \boldsymbol{M}$. This is posed with the integral

$$\iint F(\nabla \boldsymbol{J}, \boldsymbol{E}) \; dxdy, \tag{2.4}$$

where $\boldsymbol{E} = \nabla \boldsymbol{M}$ and $F(\nabla \boldsymbol{J}, \boldsymbol{E}) = \|\nabla \boldsymbol{J} - \boldsymbol{E}\|^2$. Following the formulation of Fattal et al. (2002), isolating the integrand

$$\|\nabla \boldsymbol{J} - \boldsymbol{E}\|^2 = \left(\frac{\partial \boldsymbol{J}}{\partial x} - \boldsymbol{E}_x \right)^2 - \left(\frac{\partial \boldsymbol{J}}{\partial y} - \boldsymbol{E}_y \right)^2 \tag{2.5}$$

and then by using Variational Principle, that states that a function that minimises the integral Eq. (2.4) must satisfy the Euler-Lagrange equation

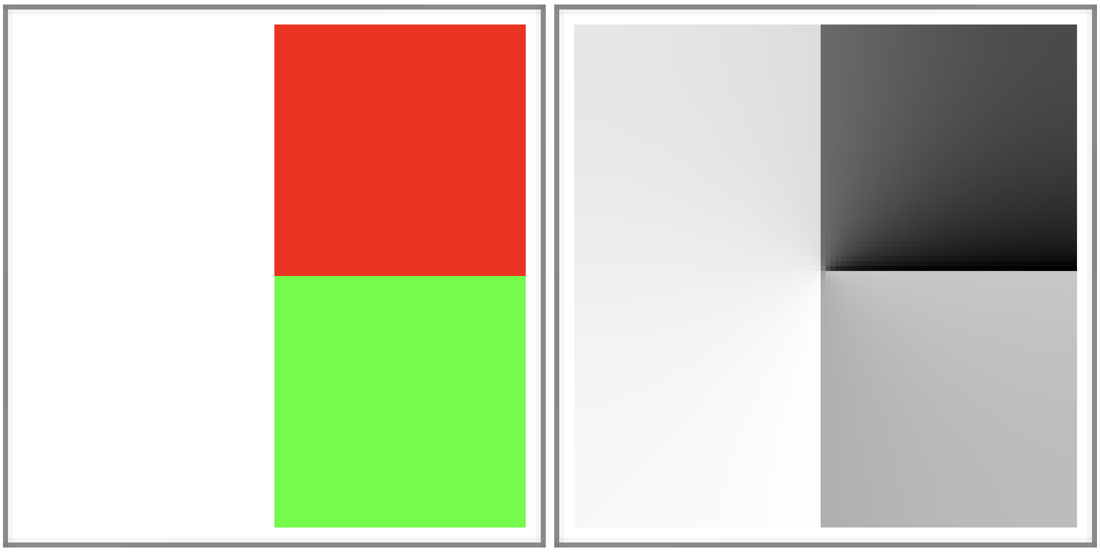
$$\frac{\partial F}{\partial J} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial \boldsymbol{J}_x} - \frac{\mathrm{d}}{\mathrm{d}y} \frac{\partial F}{\partial \boldsymbol{J}_y} = 0, \tag{2.6}$$

which is a partial differential equation in $\boldsymbol{J}$. Substituting Eq. (2.5) into Eq. (2.6) and differentiating results in the Poisson equation

$$\nabla^2 \boldsymbol{J} = \nabla \boldsymbol{E}, \tag{2.7}$$

where ∇^2 is the Laplacian operator. The task then becomes one of solving the Poisson equation where the final greyscale is found (subject to a additive constant) by integrating twice. Unfortunately $\nabla \boldsymbol{E}$ typically forms a non-conservative vector field and is non-integrable. This means our final greyscale in only a least squares approximation, which would be noticed upon viewing by the presence of smearing and halo type artefacts, see Figure. 2.1.

The challenge for researchers became that of finding the desired ‘real’ greyscale image that minimises Eq. (2.1). We shall now review the methods that have been developed to solve this problem. These methods will be split into two categories; mappings from luminance-chromaticity based colour spaces and those from the RGB colour space. In the last section we shall briefly analyse these methods to determine which one will be most suited for our multiband panchromatic image sharpening problem.



(a) Original RGB (b) Least squares solution to Poisson equation

Figure 2.1: Colour-to-greyscale example from solving Poisson's equation. Note the artefacts that can be produced in the greyscale from the optimisation. The white and grey boarders around the images serve to add clarity and were not involved in the mapping.

2.2 Luminance-Chromaticity Space Mappings

How best to incorporate the chromaticity information from a colour image when producing its greyscale has become a standard problem in the field of colour-to-grey. To simplify this process it is convenient to work in a colour space where the image can be represented by the components of visual perception: hue, saturation and brightness. Decomposing the image into its visual components allows them to be individually operated on which allows for greater control when producing a visually pleasing greyscale. There are many varieties of brightness/chromaticity based colour spaces with CIE $L^*a^*b^*$ being the most prevalent within the literature. In this colour space L^* represents the image luminance and a^* and b^* the opponent colours red-green and yellow-blue respectively. This device independent colour space has become an industry standard in colour imagery. It takes into account the non-linear response of the human visual system to optical stimuli. It is a perceptually uniform colour space, which is to say that stimuli that exhibit the same perceived differences will have the same Euclidean vector magnitudes in the colour space.

Using the resultant Euclidean magnitude between pixels as the measure of contrast to preserve, Rasche et al. (2005) posed Eq. (2.1) as a Multi-Dimensional Scaling (MDS) problem, with the constraint of preserving the luminance ordering on their final greyscale. This constraint is applied to the optimisation when the Euclidean magnitude between two pixels is lower than a threshold. In these instances an inequality is added to the optimisation and this ensures that the luminance ordering is maintained in the greyscale output. Visually this means that no contrast inversion occurs between pixels of similar chrominance. As these constraints are posed as an inequality, their solution is found within an intersection of half-spaces and solved using linear programming. To avoid the high computational cost of solving a constrained MDS problem for complex colour images, the authors proposed solving for a sparse system based on image landmark points which provides a global solution for their final greyscale by using a constrained interpolation.

Gooch et al. (2005) solved Eq. (2.1) using the gradient descent method to produce a luminance-like greyscale that has its contrast enhanced from the information contained in the chromaticity channels. Operating in the gradient domain of the luminance image, substitutions were made with the magnitude of the 2-dimensional chromatic vector gradients ($|\nabla \boldsymbol{C}|$), on the condition that it be larger. The direction of $|\nabla \boldsymbol{C}|$ was determined by what half of the polarised chrominance plane $\nabla \boldsymbol{C}$ falls within. The gradient descent method is used as their solver to produce a real image that is as close as possible to the luminance image in a least squares sense.

Neumann et al. (2007) proposed using gradient correction to address un-conservative image vector fields. Their gradient correction was conducted using a negative feedback loop, which reduces the gradient magnitudes at each iteration such that the error (which was calculated by taking the gradient sum in a square four pixel window) is below a threshold. Ideally this error would be zero for all loops, thus producing an artefact-free image. To avoid the sign problem that occurs when taking colour contrasts as the Euclidean norm, they instead used the L_3 norm metric of the (L^*, a^*, b^*) vector. They reported a much faster processing time than the iterative solution of Gooch et al. (2005).

Kuhn et al. (2008) proposed solving Eq. (2.1) using Verlet (1967)'s time-step integration on the luminance image to iteratively converge to a greyscale whose gradients are as close as can be to $\nabla \boldsymbol{M}$. To ensure that the gradient field is conservative, they apply Newtonian physics to their model as well as weights (based on the inverse of the magnitude chromaticity vector) to each luminosity pixel. This means that pixels with low saturation are permitted to change at a slower rate per iteration than high saturation ones, while pixels having zero saturation being unable to change. Due to the high computational complexity of $O(n^2)$, colour quantisation and interpolation can be applied to reduce run times.

Instead of trying to increase the contrast of the luminance image using the chromaticity information, Smith et al. (2008) produce a base greyscale L^*_{HK} from the perceptually uniform CIE $L^*u^*v^*$ colour space using the Helmholtz-Kohlrausch (HK) effect: an empirical model of assigning colour ordering for greyscale pixels that takes into account the fact that hue and saturation of two chromatically different yet isoluminant stimulii affect its perceived brightness by the human visual system. By testing various HK models, the authors decided on using Nayatani's lightness metric (Nayatani, 1997) because of its greater variance in its assignment of lightness values to pixels across the RGB spectrum. Specifically the Nayatani metric uses the images chromaticity from the perceptually uniform CIE $L^*u^*v^*$ colour space to adjust the luminance levels in L^*. Local greyscale contrasts are then enhanced in a multi-resolution framework (an adaptively-weighted multi-scale unsharp masking technique (Nowak and Baraniuk, 1998)) by modulating the bandpass levels (h_i) in a Laplacian pyramid by the root of an image consisting of the ratio

$$\lambda_i = \left(\frac{\nabla(h_i(\boldsymbol{M}))}{|h_i(L^*_{HK})|} \right)^p, \tag{2.8}$$

where $\boldsymbol{M}$ is the colour image in CIE $L^*a^*b^*$ space. Their final greyscale $\boldsymbol{J}$ is computed by superimposing the sum of the Laplacian pyramid levels upon the base greyscale

$$\boldsymbol{J} = L^*_{HK} + \sum_{i=0}^{n-1} k_i \lambda_i h_i(L^*_{HK}), \tag{2.9}$$

where parameters p ($0 \leq p \leq 1$) and k (≤ 1) exist as decision level tuning parameters that control contrast.

The unique property of using the Laplacian pyramid is that it allows for selective enhancement. Image features are segregated by size; fine details will be prominent in the lower levels whereas progressively coarser features are prominent in the higher levels. This can be advantageous when it comes to enhancing image contrast without increasing image noise.

Kim et al. (2009) optimised Eq. (2.1) by solving Poisson's equation. Their optimisation consists off finding the a^*, b^* component weights that produces a greyscale mapping that when added to L^*, minimises Eq. (2.1)

$$\boldsymbol{x} = \boldsymbol{M}_s^{-1} \boldsymbol{b}_s, \tag{2.10}$$

where $\boldsymbol{M}_s$ is a matrix that contains the gradients of the chromaticity information that is constructed from the outer-products of the ∇a^*, ∇b^* components and $\boldsymbol{b}_s$ a vector that contains the gradient enhancement (using the chromaticity components) for L^* pixels that have lost contrast from the original colour image. To prevent the computational cost of inverting a large matrix that contains all colour contrast, they allow for a limitation on the number of hues used in their optimisation. Greyscale brightness ordering from their colour gradients were assigned to that of the HK lightness predictor. If this positive or negative direction happens to be zero, the luminance direction should be used and if this fails, the sign of the L_3 norm should be adopted akin to Neumann et al. (2007).

Song et al. (2010) constructed their gradient vector field $\nabla \boldsymbol{J}$ by summing the individual gradient fields for their image over their modified hue ($\nabla \boldsymbol{H'}$), saturation ($\nabla \boldsymbol{S'}$) and lightness ($\nabla \boldsymbol{L'}$) based colour space, $\nabla \boldsymbol{J} = \frac{1}{3} \times (\nabla \boldsymbol{H'} + \nabla \boldsymbol{C'} + \nabla \boldsymbol{L'})$ and then solving Poisson's equation (Davis, 2003, W.H.Press et al., 1992, Bolz et al., 2003). The novelty of their work is how they construct these gradient vector fields; a method that they link to the psychology and biology of the human visual system. Gradients were calculated using the centre difference, with hue gradients being modulated by the square of the saturation; the rationale being that the human visual system is not sensitive to hue changes at low saturations. Adopting a similar approach to colour ordering as Gooch et al. (2005), they used a linear classifier to define a 1-dimensional space upon which to project the hue information. The gradient directions were given by the priority of hue, saturation and lastly lightness in addition to having their magnitudes modulated by the priority factors, 1, $\exp(-x_C)$ and $\exp(-x_C \exp(-x_L)$.

Du et al. (2015) proposed solving Eq. (2.1) between colour $S_{\boldsymbol{M}}(s_i)$ and greyscale $S_{\boldsymbol{J}}(s_i)$ image segments. They defined a statistical measure of colour contrast at a pixel/segment by modelling the chromaticity information with a normal distribution. Using simple linear iterative clustering (Achanta et al., 2012) to segment their colour and greyscale images, contrast/saliency values were calculated by

$$S_{\boldsymbol{M}}(s_i) = \sum_{s_i \in S} N(s_i) \exp\left(-\frac{\nabla \boldsymbol{D}_{i,j}}{\sigma^2}\right) \nabla \boldsymbol{M}_{i,j}/3 \tag{2.11}$$

and

$$S_{\boldsymbol{J}}(s_i) = \sum_{s_i \in S} N(s_i) \exp\left(-\frac{\nabla \boldsymbol{D}_{i,j}}{\sigma^2}\right) \nabla \boldsymbol{J}_{i,j}, \tag{2.12}$$

where S is the set of segments in the image, N is the number of pixels in the segment s_i and $\sigma^2 = 0.4 \times (width \times height)$, $\nabla \boldsymbol{D}_{i,j}$ is the Euclidean distance between the segment centres i and j, $\nabla \boldsymbol{M}_{i,j}$ the Euclidean distance between the segments in CIE $L^*a^*b^*$ and $\nabla \boldsymbol{J}_{i,j}$ the scalar difference between the two segments. Using the Variable-Achromatic-Colour mapping as a base greyscale L^{HK}, a series of candidate greyscale images (which is used to compute $\nabla \boldsymbol{J}$) are produced by adding on a different chromaticity mapping

$$\boldsymbol{J}_p = L^*_{HK} + k \sum_{q \in \Omega} (a_p - a_q) \exp\left(-\frac{\nabla \boldsymbol{D}_{p,q}}{\sigma^2}\right) + l \sum_{q \in \Omega} (b_p - b_q) \exp\left(-\frac{\nabla \boldsymbol{M}_{p,q}}{\sigma^2}\right), \tag{2.13}$$

where each greyscale is produced as a function of the parameters k and l. Eq. (2.1) is consequently solved by choosing the greyscale that provides the minimum value.

2.3 RGB Colour Cube Space Mappings

The sRGB colour space is the most popular among the RGB colour cube models. Invented by Microsoft and Hewlett-Packard, it was created to address the device dependency problem of the RGB model. This device dependency meant that different devices would output a different colour for a given RGB vector. This problem was exasperated with the advent of the internet when images were being viewed from different devices and displays on different browsers. The sRGB colour space was designed as a device independent consumer 'standard' colour space whose viewing characteristics matched those of the standard CRT computer monitors under typical home and office viewing conditions. As with the RGB colour model, each pixel in an image is represented by the RGB additive primary colours in the form of a 3-dimensional vector. The additive nature of these primary colours means that the pixel brightness is the sum of the vector components. Normalising each vector component at a pixel by its brightness produces a mapping to the Maxwell Triangle, which allows the saturation to be visualised as the Euclidean distance from the Triangle centre (white point - zero saturation) to its edge and the hue being the azimuth angle about said centre.

Grundland and Dodgson (2005) produced a unique means of injecting chromatic contrast into the luminance image by producing an image that contains the contrast lost from that of the colour image (in RGB space), and then adding it back onto the luminance. They first defined the contrast loss as

$$\boldsymbol{D} = 1 - \frac{\frac{1}{Y_{axis}}|\nabla \boldsymbol{Y}_{i,j}|}{\nabla \boldsymbol{M}_{i,j}}, \tag{2.14}$$

where Y_{axis} is the NTSC-Rec.601 standard luminance axis and $\nabla \boldsymbol{Y}_{i,j}$ is the luminance gradient image from their YPQ colour space. This linear luminance ($\boldsymbol{Y}$) chromaticity ($\boldsymbol{P}, \boldsymbol{Q}$) colour space can be accessed by linearly mapping the gamma corrected RGB pixels using the transform provided by the authors. Eq. (2.14) demonstrates that when the fraction is equal to 1, $\boldsymbol{D}_{i,j}$ becomes zero and this means that there has been no contrast loss from the colour image after it has been mapped to the luminance image. The opposing answer being 1 would mean all contrast has been lost. Defining the contrast loss in this way allows for it to be used as a modulating coefficient on the chromaticity gradients, whose sum defines a 2-dimensional vector of contrast loss through the chromaticity plane. The authors refer to this as the *Predominant Component* (analogous to the principal component from PCA) and it is used to produce a vector whose weights are used for reducing the dimensionality of the chromaticity components of each pixel to a single band $\boldsymbol{C}$. Their enhanced greyscale $\boldsymbol{J}$ is then produced by adding on these pixel dependent scalars to $\boldsymbol{Y}$

$$\boldsymbol{J} = \boldsymbol{Y} + \lambda \boldsymbol{C}, \tag{2.15}$$

where λ is a tuning parameter for the desired degree of contrast enhancement. Setting bounding constraints on the values each pixel can have allows for the image to take advantage of the dynamic range of the display output. Grundland and Dodgson (2005) carried this out using the saturation component $\boldsymbol{S} = \sqrt{\boldsymbol{P}^2 + \boldsymbol{Q}^2}$ and it attributes to the boundary condition if

$$y_{min} \ < \ \boldsymbol{Y} \pm \lambda \frac{y_{max}}{s_{max}} \boldsymbol{S} \ < \ y_{max}, \tag{2.16}$$

where the subscripts min/max define the dynamic range of the display. Grundland's technique also comprises further constraints which allow for greyscale preservation (grey pixels in the colour image will have the same shade in the greyscale) and luminance ordering (colour pixels of increasing luminance will be mapped to increasing grey levels).

Qiu et al. (2008) developed a constrained PCA optimisation that, subject to equality and inequality constraints produces a greyscale that possess maximum global contrast. Using global statistical variance as their metric for overall image contrast, they produced a simple optimisation based on the quadratic form

$$\max_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}, \tag{2.17}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of the colour image and $\boldsymbol{w}$ is a 3-dimensional vector that contains the RGB band weightings. To ensure that the image mean ($\mu_{\boldsymbol{I}}$) and energy are preserved, they impose two linear equality constraints on their optimisation

$$\mu_{\boldsymbol{I}} = \sum_{1}^{3} w_n \mu_n \tag{2.18}$$

and

$$1 = \sum_{1}^{3} w_n, \tag{2.19}$$

subject to the inequality constraints

$$0 \leq w_n \leq 1 : \quad n = 1, 2, 3. \tag{2.20}$$

Applying quadratic programming to maximise Eq. (2.17) outputs $\boldsymbol{w}$: a 3-dimensional vector that contains the band weights for their colour-to-grey mapping that possesses maximum global contrast.

Alsam (2009) presented a local linear colour to grey mapping. Defining contrast as the high-frequency information contained in each band, they modulate contrast lost from luminance with the colour images unsharp mask, C_{Rg}, C_{Gg}, C_{Bg} (based on a 25×25 Gaussian filter) to create an image of pixel weights

$$
\begin{aligned}
\alpha(x,y) &= \frac{C_{Rg}(x,y)}{C_{Rg}(x,y) + C_{Gg}(x,y) + C_{Bg}(x,y)} \\
\beta(x,y) &= \frac{C_{Gg}(x,y)}{C_{Rg}(x,y) + C_{Gg}(x,y) + C_{Bg}(x,y)} \\
\gamma(x,y) &= \frac{C_{Bg}(x,y)}{C_{Rg}(x,y) + C_{Gg}(x,y) + C_{Bg}(x,y)},
\end{aligned}
\tag{2.21}
$$

where the denominator normalises the pixel weights to lie between 0 and 1. Their greyscale is obtained by applying these individual pixel weights to each band and taking their sum.

Cui et al. (2010) presented a global mapping that uses the ISOMAP algorithm for their colour-to-greyscale conversion. ISOMAP was created on the basis of finding an optimum solution for maintaining the proportions between the geodesic and Euclidean magnitudes when mapping from a higher to a lower-dimensional space. This technique bears a strong resemblance to classical MDS, which Rasche et al. (2005) used in their optimisation to maintain the proportions of the Euclidean distances between their colour and greyscale images. To avoid the computational burden of calculating all distances between pixels, the authors proposed using the landmark ISOMAP algorithm that finds sparse nodes in a kd-tree. This essentially means that their landmarks are based on outliers that are defined by a user controlled parameter. Additional contrast enhancement of their final greyscale was made available by allowing for the manipulation (using nonlinear scaling x^{λ}) of the geodesic distances. By treating each colour vector as the node on a graph, a curved manifold is created. The structure of the manifold depends on determining which nodes belong to which neighbourhood, solved by a user parameter k that states how many neighbours a node can have. Computing the Euclidean magnitude between the k-nearest neighbour nodes defines the graph edge weights. Summing the edge weights between two nodes of interest allows for the shortest path to be found, i.e. the geodesic distance, which when applied to the entire graph results in a matrix of geodesic distances. Applying classical MDS to this matrix outputs a lower-dimensional representation that preserves the original manifolds intrinsic geometry.

Lu et al. (2012a) solved Eq. (2.1) using the fixed-point iteration method to produce a globally mapped greyscale that has been constructed from a weighted linear, bilinear and quadratic combination of the RGB monomials. The bulk of their contribution was in designating a sign

$\oplus$ to their colour ordering of $\nabla \boldsymbol{M}$. This was computed by using the signs that maximise the likelihood under the assumption that the colour differences follow a bimodal distribution. As maximising the likelihood is the same as minimising its negative logarithm

$$E(\boldsymbol{w}) = -\sum_{(i,j)\in\mathcal{N}} \ln\left\{\alpha_{i,j}\exp\left\{-\frac{|\nabla \boldsymbol{J}_{i,j} - \nabla \boldsymbol{M}_{i,j}|^2}{2\sigma^2}\right\} + (1-\alpha_{i,j})\exp\left\{-\frac{|\nabla \boldsymbol{J}_{i,j} + \nabla \boldsymbol{M}_{i,j}|^2}{2\sigma^2}\right\}\right\}, \tag{2.22}$$

where $\mathcal{N}$ is a four neighbour set, $\boldsymbol{w}$ is a 9-dimensional vector whose components contain the monomial weights and

$$\alpha_{i,j} = \begin{cases} 1.0 & \text{if} \quad \boldsymbol{R}_i \leq \boldsymbol{R}_j, \boldsymbol{G}_i \leq \boldsymbol{G}_j, \boldsymbol{B}_i \leq \boldsymbol{B}_j \\ 0.5 & \text{otherwise.} \end{cases} \tag{2.23}$$

In a second paper Lu et al. (2012b) produced a means of reducing the time complexity of their technique so that it can be used in real-time video. Uniformly sampling band weighting space defined by Eq.(2.19) and (2.20), they produce a library of greyscales (by linearly mapping), with their final greyscale being chosen on the basis that it minimises Eq. (2.22).

This process of creating a library of greyscales and choosing the one that satisfies our desired properties is a Look Up Table (LUT) approach to the problem. Song et al. (2013) have also utilised this method with the objective of preserving global contrast in their greyscale, which they find by finding the minimum of Eq. (2.1) between bilaterally filtered (Yang et al., 2009) greyscale images ($\boldsymbol{J}^{\boldsymbol{\Upsilon}}$) and the original colour RGB ($\boldsymbol{M}^{\boldsymbol{\Upsilon}}$).

Bilateral filters are edge preserving operators defined as

$$\boldsymbol{M}^{\boldsymbol{\Upsilon}}(p) = \frac{\sum_{\boldsymbol{q}\in\Omega_{\boldsymbol{p}}} G_{\sigma_s}(||\boldsymbol{p}-\boldsymbol{q}||)G_{\sigma_r}(||\boldsymbol{\Upsilon}(\boldsymbol{p}) - \boldsymbol{\Upsilon}(\boldsymbol{q})||)\boldsymbol{M}(\boldsymbol{q})}{\sum_{\boldsymbol{q}\in\Omega_{\boldsymbol{p}}} G_{\sigma_s}(||\boldsymbol{p}-\boldsymbol{q}||)G_{\sigma_r}(||\boldsymbol{\Upsilon}(\boldsymbol{p}) - \boldsymbol{\Upsilon}(\boldsymbol{q})||)}, \tag{2.24}$$

where $\boldsymbol{q}$ is a pixel in the neighbourhood $\Omega_{\boldsymbol{p}}$ of pixel $\boldsymbol{p}$, and G_{σ_s} and G_{σ_r} are the spatial and range filter kernels measuring the spatial and colour/intensity similarity, and $\boldsymbol{\Upsilon}$ is the guidance image, which in this case is the RGB image ($\boldsymbol{M}$) and a series of linearly mapped greyscales ($\boldsymbol{J}$) constructed from the weights in Eq. (2.19).

They also extended the work to real-time (Song et al., 2014). Again with the use of bilateral filtering, they estimated the contrast lost in the luminance mapping by creating a 3-dimensional contrast loss image

$$\boldsymbol{D} = \boldsymbol{M}^{\boldsymbol{M}} - \boldsymbol{M}^{\boldsymbol{L}}, \tag{2.25}$$

where the colour image that has been bilaterally filtered with both itself ($\boldsymbol{M}^{\boldsymbol{M}}$) and the luminance ($\boldsymbol{M}^{\boldsymbol{L}}$) image respectively and $\boldsymbol{D}$ is the 3-dimensional contrast loss image. The authors proposed using this contrast loss image to identify which vector component at a pixel to use from the gradient colour image to represent the desired contrast in the greyscale. That is to say, at a pixel, the largest vector component in $\boldsymbol{D}$ defines the vector component from the colour gradient image to use as the metric of colour contrast to be represented in the greyscale, i.e. $\nabla\boldsymbol{M}_{x,y}$ from Eq. (2.1). The objective then becomes one of solving for band weights $\boldsymbol{w}$ that can provide a linear mapping of $\boldsymbol{D}$ that can be mapped to the luminance image $\boldsymbol{L}$ that minimises Eq. (2.1) i.e. $\nabla\boldsymbol{D}\boldsymbol{w} + \nabla\boldsymbol{L} = \nabla\boldsymbol{J}_{x,y}$.

2.4 Analysis of Colour-to-Greyscale Linear Mappings

We seek to choose a method of converting a colour image to a greyscale, that would be applicable to multispectral (> 3 bands) imagery and the pan-sharpening problem. Most of the early methods in colour-to-grey operate in the perceptually uniform CIE $L^*a^*b^*$ colour space: a 3-dimensional space that represents the image by L^* the lightness, a^* and b^* the colour opponents green-red and blue-yellow respectively. From these chromaticity components, the hue and saturation can be computed showing the resemblance to the standard IHS space commonly used in pan-sharpening. Consequently, CIE $L^*a^*b^*$ shares the same shortcomings of the IHS method in that it requires a transformation from RGB space meaning that it is limited to three bands.

None of the methods mentioned minimise Eq. (2.1) in closed form and require the use of iterative search based algorithms to produce their solutions. This increases the complexity of the method.

Methods that use the RGB colour cube model offer the advantages that the inclusion of the $\boldsymbol{NIR}$ band comes as a natural addition to the dimension of the vector space. Furthermore, they also encompass the methods that have low complexity, the simplest being that of globally projecting the RGB bands onto a vector. Provided that the components of this vector are all greater than zero, image brightness ordering is maintained as a consequence of the additive nature of the colour cube model. The challenge of these methods becomes one of finding the vector that satisfies the optimisation with respect to any imposed constraints.

Using their predominant component analysis, Grundland and Dodgson (2005) define a direction of contrast loss in the chrominance plane with the standard luminance mapping. Projecting the chrominance information onto this vector and then adding it onto the luminance increases the contrast in the greyscale, which is further enhanced when the saturation component is used to increase the dynamic range of their greyscale. Although this method operates in their own luminance/opponent based colour space, the colour image contrast is defined by the Euclidean magnitude of the RGB vectors which can extend to inclusion of the $\boldsymbol{NIR}$ band. The disadvantage is that they rely on the gradient image of the luminance mapping to define the direction of the colour image contrast, which does not take into consideration the $\boldsymbol{NIR}$ band.

Qiu et al. (2008) posed a constrained PCA based method that produces a projection vector that maximises the image variance subject to preserving the statistical mean of the colour image. This optimisation is designed to output maximum global greyscale image contrast and can be easily extended for use with multispectral imagery. It has advantages over LUT approaches in that it produces a mathematically optimum answer to the problem. The disadvantage being that it is currently solved using the complex quadratic programming algorithm.

Alsam's (Alsam, 2009) local approach computes a projection vector for each pixel in the colour image. This approach requires creation of a high-frequency edge map of the RGB bands and a low-frequency map of the luminance image to calculate individual pixel weights to reinforce edges in the luminance image. Creation of these maps not only increases complexity but also involves user defined parameters that would need a trial and error style testing to produce the satisfactory results. A further disadvantage to this technique is that it would increase any sensor noise present in the image.

Song et al. (2013) creates a LUT of projection vectors. For each colour image, they produce a library of band weights by uniformly sampling coefficient space and then selecting the greyscale that has a contrast (defined by joint bilateral filter) that is as close as possible to the colour input. The use of the linear energy preserving constraint of Qiu et al. (2008) also means the method can be constrained further to preserve image brightness. Additional advantages to the

LUT approach is in its flexibility with respect to its objective function. The method falls short in that the greyscale output is only a close approximation to the optimal greyscale; a consequence of the discrete selection of band weights in the LUT. The authors argue that this approximation would look identical to a human observer to that of the optimum greyscale, however this would also mean that its scores would be lower than the optimum when evaluated using numerical metrics. Applying this method to the IHS pan-sharpening problem requires producing a LUT that includes the $\boldsymbol{NIR}$ component, and then selecting the image that has the highest correlation with that of $\boldsymbol{P}$. As this is a simple extension to their work, we feel there is not enough scope for novelty to warrant further development.

2.5 Conclusion

In this chapter we have reviewed the state of the art in colour-to-greyscale mappings, where the objective is to replicate colour contrasts in its greyscale mapping. Typically this is carried out by either injecting the chromaticity information into the luminance mapping from a perceptually uniform colour space, or from linearly mapping the colour bands from the additive RGB colour cube.

From this review we feel that the work of Qiu et al. (2008) has the greatest synergy to the pan-sharpening problem. The optimisation is based on linearly mapping from the RGB colour space which can be naturally extended to multispectral imagery. From the construction of their model, linear constraints can be easily applied as has been demonstrated by the authors with regard to their constraints on preserving image energy and statistical mean. Although the optimisation currently uses a quadratic programming algorithm to solve, we shall show in the next chapter that it can be solved in closed form, thus reducing its computational complexity.

Chapter 3

Colour-to-Greyscale: Linear Mappings

So far we have described how producing a single band image from a linear mapping of multiple bands has applications in panchromatic image sharpening and in constructing a visually pleasing greyscale image. The problem behind both applications is how to compute the band weights such that high-frequency information is preserved in the greyscale. Qiu et al. (2008) designed a colour-to-greyscale algorithm that maximises contrast in the greyscale with an optimisation that we recognise as being fundamental to both colour-to-grey and pan-sharpening image fusion. We start this chapter by reviewing their optimisation and exploring in detail the objective function posed by Qiu et al. (2008). Currently solved using quadratic programming, we shall show how their optimisation can be solved using standard arithmetic calculations of the image band means. In addition to this, we further their work by showing how a simple modification to their optimisation can produce band weights that produces full dynamic range greyscale imagery by number sorting.

3.1 Linear Mapping

Linear mappings offer the lowest time complexity for producing a greyscale from a colour image. In Chapter. 1 we defined linearly mapping of image bands to be

$$\boldsymbol{J} = \sum_{1}^{d} w_n \boldsymbol{M}_n, \tag{3.1}$$

where $\boldsymbol{J}$ is the single band greyscale, d is the total number of bands being used in the fusion, w_n is the n^{th} band weight and $\boldsymbol{M}_n$ is the n^{th} multispectral band.

The standard choice of band weights for colour-to-greyscale is to map the RGB pixel triplet to a single intensity value. As we described in Chapter 1, this is achieved by setting the band weights to 1/3

$$\boldsymbol{I} = \frac{\boldsymbol{R} + \boldsymbol{G} + \boldsymbol{B}}{3}. \tag{3.2}$$

The numerator of Eq. (3.2) defines the arithmetic sum for the RGB triplet at a pixel and we refer to this as its *brightness*. It is clear that by mapping pixels with different RGB components yet identical brightness values will result in pixels being mapped to the same shade of grey, and ultimately a decrease in global image contrast and information. An image dependent means of determining the band weights must be used so that these instances of contrast loss in the greyscale can be minimised.

These weights are generally constrained to sum to unity

$$1 = \sum_{1}^{3} w_n, \tag{3.3}$$

which preserves (i) the spectral contribution and (ii) the greyscale values already present in the RGB. In addition to Eq. (3.3), the boundary conditions

$$0 \leq w_n \leq 1 : \quad n = 1, 2, 3 \tag{3.4}$$

ensure that no clipping occurs in the greyscale.

Compared to the intensity mapping (Eq. (3.2)), the colour-to-greyscale methods reviewed in

Chapter 2 enhance the contrast of the greyscale reproduction at a cost of altering its brightness (which is preserved in the $\boldsymbol{I}$ mapping). With these observations in mind, Qiu et al. (2008) novelly considered maximising contrast while simultaneously maintaining the reproduction brightness: the brightness preserving greyscale conversion.

3.2 Brightness Preserving Greyscale Conversion: Preliminaries

Using global statistical variance as their metric for overall image contrast, Qiu et al. (2008) produced an optimisation based on the quadratic form

$$\max_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}, \tag{3.5}$$

where the covariance matrix $\boldsymbol{\Sigma}$ of the colour image is computed by

$$\Sigma_{m,n} = E[(m_m - \mu_m)(m_n - \mu_n)] \qquad \forall \quad m,n \in \{\boldsymbol{R}, \boldsymbol{G}, \boldsymbol{B}\}. \tag{3.6}$$

Here E is the expectation operator, m_n is a pixel value in a given band n and μ_n is the mean of that band. To ensure that the image brightness is preserved and that all output greyscale pixels lie between 0 and 1, they impose the equality constraints

$$\boldsymbol{w}^T \boldsymbol{\mu} = \boldsymbol{u}^T \frac{\boldsymbol{\mu}}{3} \tag{3.7}$$

and

$$\boldsymbol{w}^T \boldsymbol{u} = 1, \tag{3.8}$$

in addition to the inequality constraints in Eq. (3.4). Here $\boldsymbol{\mu}$ is a vector whose components contain the statistical mean of the RGB bands. In Eq. (3.8) we have re-written Eq. (3.3) as a dot product between the band weighting vector $\boldsymbol{w}$ and the unitary vector $\boldsymbol{u}$

$$\boldsymbol{w}^T = [\alpha \quad \beta \quad \gamma], \quad \boldsymbol{u}^T = [1 \quad 1 \quad 1] \quad \text{and} \quad \boldsymbol{\mu}^T = [\mu_R \quad \mu_G \quad \mu_B]. \tag{3.9}$$

Eq. (3.5) represents the quadratic objective function which is satisfied when $\boldsymbol{w}$ provides a maximum variance subject to the constraints in Eqs. (3.8), (3.7) and the boundary condition Eq. (3.4).

As the covariance matrix is positive-definite, the calculation of variance using the quadratic form is always positive

$$\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} > 0 \quad \forall \quad \boldsymbol{w} \in \mathbb{R}^n. \tag{3.10}$$

As with all positive-definite quadratic equations, there exists a single root. By exploiting a property of the positive-definite matrix, the quadratic form posed by Qiu et al. (2008) can be reduced to a vector magnitude.

By computing the square root of $\boldsymbol{\Sigma}$ (Golub and Loan, 1996) by singular value decomposition (SVD)

$$\sqrt{\boldsymbol{\Sigma}} = \boldsymbol{U} \sqrt{\boldsymbol{S}} \boldsymbol{U}^T \tag{3.11}$$

and

$$\boldsymbol{\Sigma} = \sqrt{\boldsymbol{\Sigma}} \sqrt{\boldsymbol{\Sigma}}. \tag{3.12}$$

Substituting Eq. (3.12) into the quadratic form that defines variance (Finlayson and Matheson, 2012) gives

$$\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} = \boldsymbol{w}^T \sqrt{\boldsymbol{\Sigma}} \sqrt{\boldsymbol{\Sigma}} \boldsymbol{w} = \dot{\boldsymbol{w}}^{\mathrm{T}} \boldsymbol{I} \dot{\boldsymbol{w}} = \|\dot{\boldsymbol{w}}\|_2^2, \tag{3.13}$$

where $\boldsymbol{I}$ is the identity matrix and $\dot{\boldsymbol{w}} = \sqrt{\boldsymbol{\Sigma}} \boldsymbol{w}$. $\sqrt{\boldsymbol{\Sigma}}$ linearly transforms the band weighting vector $\boldsymbol{w}$ in such a way that the quadratic form is reduced to a simple vector magnitude of the vector $\dot{\boldsymbol{w}}$.

3.2.1 The Positive-Definite Quadratic Form and Convex Sets

Linear equality constraints define hyper-planes that are sets of the form (Boyd and Vandenberghe, 2009)

$$\left\{ \boldsymbol{x} | \boldsymbol{a}^T \boldsymbol{x} = b \right\}, \tag{3.14}$$

where $\boldsymbol{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$. $\boldsymbol{a}$ is the normal vector to the hyper-plane whose components are defined by coefficients of the variables $\boldsymbol{x}$. Each plane divides $\mathbb{R}^n$ into two half-spaces. A half-space is a set of the form

$$\left\{\boldsymbol{x}|\boldsymbol{a}^T\boldsymbol{x} \leq b\right\}, \tag{3.15}$$

where the boundary of a half-space is described by the equation of a plane (Eq. (3.14)). The intersection of closed half-spaces describe a convex set. A known property of convex sets is that any vector that lies within the convex hull of its vertices $\boldsymbol{v}$ can be obtained by their convex combination

$$\boldsymbol{c} = \sum_{i=1}^{n} x_i\boldsymbol{v}_i, \qquad \text{s.t} \quad x_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{n} x_i = 1, \tag{3.16}$$

where x_i represents the weight of the i^{th} vertex $\boldsymbol{v}_i$. We can write this in the form $\boldsymbol{Ax} = \boldsymbol{c}$ where the columns of $\boldsymbol{A}$ contain the vectors that represents a vertex, $\boldsymbol{x}$ the vector containing our vertex weights and $\boldsymbol{c}$ the vector that defines our interior point. Substituting into our quadratic form (Eq. (3.5))

$$\max_{\boldsymbol{x}} \quad \boldsymbol{x}^T\boldsymbol{V}\boldsymbol{x}, \quad \boldsymbol{V} = \boldsymbol{A}^T\sqrt{\boldsymbol{\Sigma}}\sqrt{\boldsymbol{\Sigma}}\boldsymbol{A}, \tag{3.17}$$

where $\boldsymbol{V}$ is a positive-definite matrix that has been formed by the inner cross product of the matrix $\boldsymbol{A}\sqrt{\boldsymbol{\Sigma}}$. This new quadratic form is a function of the vertex weights that are subject to the constraints in Eq. (3.16). Reducing this new quadratic form down to a vector magnitude by making the substitution $\dot{\boldsymbol{x}} = \sqrt{\boldsymbol{V}}\boldsymbol{x}$ gives

$$\|\dot{\boldsymbol{x}}\|_2 = \|\dot{\boldsymbol{c}}\|_2 > 0. \tag{3.18}$$

$\|\dot{\boldsymbol{x}}\|_2$ can be increased by scaling $\dot{\boldsymbol{x}}$ with a number greater than 1. As linear mappings are operation preserving transforms, this scaling would be equivalent to increasing the magnitude of the vector $\boldsymbol{x}$. From this we can say that in maximising $\|\boldsymbol{x}\|_2$ we in turn maximise our variance $\|\dot{\boldsymbol{x}}\|_2$. Recall that the constraints on our vertex weights sums to unity: $\|\boldsymbol{x}\|_1 = 1$ and this describes a plane under cubic constraints that is identical to Eq. (3.8) which visually is the $(+ +\dots +)$ quadrant of a L_1 norm-ball. We therefore wish to find the $\boldsymbol{x}$ that has maximum Euclidean magnitude subject to $\|\boldsymbol{x}\|_1 = 1$.

The relationship between the L_2 norm (Euclidean magnitude) to each point on the normalised

L_1 and L_2 norm-ball can be defined as

$$\left|\left|\frac{\boldsymbol{x}}{||\boldsymbol{x}||_1}\right|\right|_2 \leq \left|\left|\frac{\boldsymbol{x}}{||\boldsymbol{x}||_2}\right|\right|_2 = 1. \tag{3.19}$$

By using *Hölders* (and its special case, the *Cauchy-Schwarz*) inequality, the well known inequality relationship between L_1 and L_2 norms can be obtained

$$||\boldsymbol{x}||_2 \leq ||\boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i| = \sum_{i=1}^{n} |\boldsymbol{x}_i| \cdot 1 \leq \left(\sum_{i=1}^{n} |\boldsymbol{x}_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} 1_i^2\right)^{\frac{1}{2}} = \sqrt{n}\,||\boldsymbol{x}||_2\,, \tag{3.20}$$

where n is the dimensionality of our vector. From Eq. (3.20) $||\boldsymbol{x}||_2 \leq ||\boldsymbol{x}||_1 \leq \sqrt{n}||\boldsymbol{x}||_2$, the denominator on the left side of the inequality in Eq. (3.19) will indeed be larger than or equal to the denominator on the right side (in agreement with Eq. (3.19)). This result implies that the maximum L_2 of our L_1 magnitude will exist at the upper bound of the left side of the inequality of Eq. (3.19), i.e. where $||\boldsymbol{x}||_1 = ||\boldsymbol{x}||_2$. The only time when this can occur is when $||\boldsymbol{x}||_1 = \sqrt{n}||\boldsymbol{x}||_2$ which only happens when n is equal to 1. That is to say the L_1 and L_2 metric equal each other when the dimensionality of our vector is 1 which when considering Eq. (3.16) and that $0 \leq x_i \leq 1\ \forall\ i \in \boldsymbol{x}$ is true *iff* $\boldsymbol{x}_i \in \mathbb{R}^n = 1$.

This general result tells us that the vector that maximises our quadratic form is located where $x_i = 1$; a single vertex $\boldsymbol{v}$ of our convex polytope (which could be of any size and possess any number of vertices). All that is left for us to do, is check each vertex and select the one that has the greatest measure of variance. As we have shown, this is true for all positive-definite quadratic forms (as it is independent of the images covariance matrix) it is invariant of the image. In the next section we shall describe how to compute the convex polytope from an intersection of half-spaces.

3.2.2 Intersection of Half-spaces

The solution to the intersection of half-spaces problem posed by Preparata (1985) will be described in this section. Although the solution applies to half-spaces of any dimension, for ease of understanding we refer to 2-dimensional and 3-dimensional cases. To begin, we introduce homogeneous coordinates (C.R. Wylie, 2008).

Consider a 2-dimensional vector space $\mathbb{R}^2$. If (x, y) are the rectangular coordinates, of an arbitrary point, $\boldsymbol{p}$, and if (x_1, x_2, x_3) are any real numbers such that $x_1/x_3 = x$ and $x_2/x_3 = y$, then the triple x_1, x_2, x_3 is said to be a set of homogeneous coordinates for $\boldsymbol{p}$.

The original rectangular coordinates (dimensionality $d = 2$) (x, y) are often referred to as *nonhomogeneous coordinates.* The homogeneous coordinate representation adds an extra dimensionality $(d + 1)$, x_3, to our vector space, and that the rectangular coordinate $(x, y, z) = (x_1/x_3, x_2/x_3, x_3/x_3)$ is simply a *central projection* of the homogeneous coordinate to the (projected) plane $z = 1$, that contains our rectangular coordinate vector space. It is clear from the description of the homogeneous coordinate, that the unique instance when $x_3 = 0$ defines points at infinity in our projected plane/rectangular coordinate space, and lies within the subspace defined by the plane $z = 0$ in $\mathbb{R}^3$.

The homogeneous coordinate representation of a point also extends to geometric structures such as lines, hyperplanes and conics. For example, the equation of a line

$$ax + by + c = 0, \tag{3.21}$$

becomes that of a plane that passes through the origin of $\mathbb{R}^3$ in homogeneous coordinates

$$ax_1 + bx_2 + cx_3 = 0. \tag{3.22}$$

By setting $x_3 = 1$ returns us to our original equation of a line. Additionally, the instances when $x_3 = 0$ defines the equations when the lines pass through the origin of $\mathbb{R}^2$.

Equations of lines can be expressed with a homogeneous coordinate representation that is determined by their coefficients $[a, b, c]$ (C.R. Wylie, 2008). If $ax + by + c = 0$ is the equation of a line, p, in rectangular coordinates, then the triple $[a, b, c]$ is said to be a set of homogeneous coordinates for p.

To distinguish between triples that are coordinates of a line and those that are coordinates of a point, we shall adopt the notation of C.R. Wylie (2008) who uses square brackets for the former and parentheses for the latter. If $[a, b, c]$ are the coordinates of a line p, then the equation of p in homogeneous point-coordinates is $ax_1 + bx_2 + cx_3 = 0$. If (x_1, x_2, x_3) are the homogeneous coordinates of a point $\boldsymbol{p}$ then the equation of $\boldsymbol{p}$ in line-coordinates is $x_1 a + x_2 b + x_3 c = 0$.

There exists a relationship between the Euclidean distance d_l of a line in $\mathbb{R}^2$ and its non-homogeneous point coordinates. The Euclidean distance of a line from the origin of $\mathbb{R}^2$ is

$$d_l = \frac{|c|}{\sqrt{a^2 + b^2}}. \tag{3.23}$$

If we substitute Eq. (3.23) into Eq. (3.21) and calculate the Euclidean distance d_h of its homogeneous coordinates from the origin of $\mathbb{R}^2$, it cancels down to

$$d_h = \frac{1}{d_l}, \tag{3.24}$$

which simply states that the distance of a line from the origin is inversely proportional to the distance of its homogeneous point coordinates. If we imagine that the equation of a line is tangent to a unit circle in $\mathbb{R}^2$, then its homogeneous coordinates will define a point that lies on the exact opposite point of the circle, and this is true for all tangent lines and their homogeneous coordinate points. This scenario portrays another property of this transform of lines to points, and that is it preserves the *incidence ordering.* If we were to look at the ordering of lines relative to each with respect to the unit circle, then their homogeneous points would possess the same order.

This process of mapping a line into a point, or a point into a line is known as a *duality transform*, and objects that have been mapped to its dual are said to reside in its dual (as opposed to its primal) space. There are many different types of duality transforms and this one we have described is known as the duality with respect to the unit circle (Franco P. Preparata, 1985).

Intersections of many closed half-spaces will often define a set which only a portion have contributed to the boundary/faces of the polytope. The remaining half-spaces are *redundant.* Provided that the origin of our coordinate system lies within the intersection (and if it doesn't, we can simply translate the half-spaces until it is) of half-spaces, we can apply the duality transform with respect to a unit circle to exploit the inverse proportionality property between the distances of the lines and their homogeneous coordinates. The half-spaces that we contribute to the bounded solution will be ones whose homogeneous coordinates contribute to the convex hull that encloses the homogeneous points for all the half-spaces.

Algorithm 1 Half-space Intersection.

```
1: procedure IntersectHalf-spaces
2:     Input: Array of half-spaces $\Pi_i$ in form $\boldsymbol{a}_i^T \boldsymbol{x} \leq b_i$
3:     Output: Vertices of the convex polytope ($\boldsymbol{v}_k$) defined by the half-space intersections
4:     $\Pi^+ \leftarrow$ duality transform $\Pi$ with respect to the unit circle
5:     $\boldsymbol{V} \leftarrow \mathrm{conv}(\Pi^+)$                    ▷ Removes redundant half-spaces
6:     $\boldsymbol{V} \leftarrow$ remove colinear points                  ▷ Vertices in dual space
7:     $\boldsymbol{u} \leftarrow \dim(\boldsymbol{x})$-dimensional vector of minus ones
8:     for $k$ from 1 to length of $\boldsymbol{V}$ do
9:         $\boldsymbol{v}_k \leftarrow \boldsymbol{V}(k)^{-1}\boldsymbol{u}$   ▷ Solving for the vertices in the primal space
10:    Return $\boldsymbol{v}_k$
```

Algorithm. 1 summarises the steps for computing the vertices that define the polytope from an intersection of half-spaces. Here we use the duality transform with respect to the unit circle (line 4). After the redundant half-spaces have been removed (line 5) we remove colinear dual points as these represent parallel half-spaces which will never intersect. The final step (line 9) is carried out by computing the vertices in the primal space by computing the intersection of the ordered polytope edges by simply solving a set of linear equations for each vertex.

3.3 Brightness Preserving Colour-to-Greyscale

We start by labelling Eqs. (3.8), (3.7) and the boundaries of the three half-spaces from Eq. (3.4) that intersect with the origin as linear subsets in $\mathbb{R}^3$

$$E = \{\boldsymbol{w} \in \mathbb{R}^3 : \boldsymbol{w}^T\boldsymbol{u} = 1\} \subseteq \mathbb{R}^3, \tag{3.25}$$

$$B = \{\boldsymbol{w} \in \mathbb{R}^3 : \boldsymbol{w}^T\boldsymbol{\mu} = \boldsymbol{u}^T\frac{\boldsymbol{\mu}}{3}\} \subseteq \mathbb{R}^3, \tag{3.26}$$

$$C_i = \{\boldsymbol{w} \in \mathbb{R}^3 : \boldsymbol{w}^T\boldsymbol{e}_i = 0 \quad | \quad i = 1, 2, 3\} \subseteq \mathbb{R}^3, \tag{3.27}$$

where

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.28}$$

$L = B \cap E$ describes a line in $\mathbb{R}^3$. The boundaries of the half-spaces (Eq. (3.27)) truncate

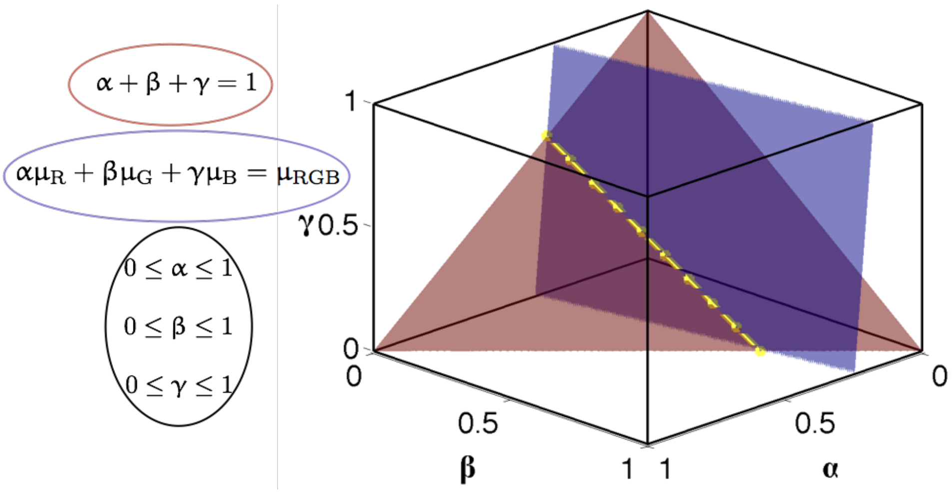


Figure 3.1: Geometric representation of the closed form solution. Here the sets have been assigned a colour: E, B and C ($\alpha = \beta = \gamma = 0$). The dashed yellow line represents the intersection of the sets E and B to give the solution set L.

this line to the convex set S (see Figure. 3.1)

$$S = \{\rho \in \mathbb{R} : \rho \boldsymbol{w}_x + (1-\rho)\boldsymbol{w}_y \quad | \quad 0 \leq \rho \leq 1\} \subseteq L \tag{3.29}$$

where $\boldsymbol{w}_{x|y} \in S$ define the two vertices of the line segment and ρ is the mixing coefficient. As we know from Section 3.2.1, the vertices of our convex set will be candidates for maximum variance. We can readily solve for our vertices by solving for $E \cap B \cap C_i$ using the tools of linear algebra. A generalised solution can be formulated using only the RGB band means which we shall now describe.

The set of all vectors $\boldsymbol{w} \in \mathbb{R}^3$ that lie in both B and E are given by their intersection L. For $\boldsymbol{w} \in L$ we define the variance K by the quadratic form Eq. (3.5)

$$K : L \to \mathbb{R}; \boldsymbol{w} \mapsto \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}. \tag{3.30}$$

The set L is a line in $\mathbb{R}^3$

$$\boldsymbol{w} = k\boldsymbol{n} + \boldsymbol{i}, \tag{3.31}$$

where k is our variable in $\mathbb{R}$ and $\boldsymbol{i}$ is our known vector in $\mathbb{R}^3$ that defines a point on the line, whose components are the band weights for the standard intensity mapping

$$\boldsymbol{i}^T = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \tag{3.32}$$

and $\boldsymbol{n}$ is the basis which provides the direction of our line L. This can be computed by a cross product between the normal vectors of Eqs. (3.25) and (3.26)

$$\boldsymbol{N} = \text{Null}\left(\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{\mu} \end{bmatrix} \right) = \boldsymbol{n} = \boldsymbol{u} \times \boldsymbol{\mu} = \begin{bmatrix} \mu_B - \mu_G \\ \mu_R - \mu_B \\ \mu_G - \mu_R \end{bmatrix}. \tag{3.33}$$

If we substitute Eq. (3.31) into Eq. (3.27)

$$k(\boldsymbol{n}^T \cdot \boldsymbol{e}_i) + \frac{1}{3}(\boldsymbol{u}^T \cdot \boldsymbol{e}_i) = 0 \tag{3.34}$$

and solve for k

$$k = -\frac{1}{3(\boldsymbol{n}^T \cdot \boldsymbol{e}_i)} \tag{3.35}$$

allows for us to calculate $\boldsymbol{w}_i$ by making the substitution of k into our equation of a line Eq. (3.31)

$$\boldsymbol{w}_i = \frac{1}{3}\left[-\frac{\boldsymbol{n}}{(\boldsymbol{n} \cdot \boldsymbol{e}_i)} + \boldsymbol{u} \right]. \tag{3.36}$$

Finally by substituting our vectors defined by Eqs. (3.28) and (3.33) into Eq. (3.36) gives us the three intersections at the half-space boundaries $\alpha = \beta = \gamma = 0$

$$\boldsymbol{w}_1 = \frac{1}{3}\begin{bmatrix} 0 \\ \frac{\mu_B - \mu_R}{\mu_B - \mu_G} + 1 \\ \frac{\mu_R - \mu_G}{\mu_B - \mu_G} + 1 \end{bmatrix}, \quad \boldsymbol{w}_2 = \frac{1}{3}\begin{bmatrix} \frac{\mu_G - \mu_B}{\mu_R - \mu_B} + 1 \\ 0 \\ \frac{\mu_R - \mu_G}{\mu_R - \mu_B} + 1 \end{bmatrix}, \quad \boldsymbol{w}_3 = \frac{1}{3}\begin{bmatrix} \frac{\mu_R - \mu_G}{\mu_G - \mu_R} + 1 \\ \frac{\mu_B - \mu_R}{\mu_G - \mu_R} + 1 \\ 0 \end{bmatrix} \tag{3.37}$$

which can be re-written as

$$\boldsymbol{w}_1 = \frac{1}{3(\mu_B - \mu_G)} \begin{bmatrix} 0 \\ 2\mu_B - \mu_R - \mu_G \\ \mu_R + \mu_B - 2\mu_G \end{bmatrix},$$

$$\boldsymbol{w}_2 = \frac{1}{3(\mu_R - \mu_B)} \begin{bmatrix} \mu_G + \mu_R - 2\mu_B \\ 0 \\ 2\mu_R - \mu_B - \mu_G \end{bmatrix}, \tag{3.38}$$

$$\boldsymbol{w}_3 = \frac{1}{3(\mu_G - \mu_R)} \begin{bmatrix} 2\mu_G - \mu_B - \mu_R \\ \mu_B + \mu_G - 2\mu_R \\ 0 \end{bmatrix}.$$

One of the three vectors in Eq. (3.38) will contain a negative component. As this vector defines an intersection that violates the cubic boundary condition we can dismiss it, leaving us with two vectors that are the end points ($\boldsymbol{w}_x$ and $\boldsymbol{w}_y$) of S. Substituting each of these two vectors into the quadratic form Eq. (3.5) will give us two calculations of variance that our greyscale will possess should we use them in our linear mapping. Choosing the vector that produces the largest variance is the weighting vector that we seek for our closed form maximum variance, brightness preserving colour-to-greyscale solution. As this closed form solution produces the same greyscale as that of Qiu et al. (2008), we shall call refer to it as $\boldsymbol{Q}$.

The pseudocode is shown in Algorithm. 2, where lines 8 to 12 represent the solution we have just presented. Lines 4 to 6 represent the solution to the degenerate case which we will now discuss.

3.3.1 Degenerate Case

Eq. (3.38) define three vectors that are computed by arithmetic operations on the band means of our colour image. So far we have assumed (fairly) that these band means are all different. There exists an unlikely scenario where they are identical (for example, if our image had been

Algorithm 2 Brightness Preserving Colour to Grey.

```
 1: procedure BRIGHTNESSPRESERVINGCOLOUR2GREY
 2:     Input: Colour image, $\boldsymbol{M}$
 3:     Output: Image band weights $\boldsymbol{w}$ and greyscale $\boldsymbol{Q}$
 4:     if $\mu_R = \mu_G = \mu_B$ then                                  ▷ Degenerate case
 5:         $\boldsymbol{w}_1 = [1 \quad 0 \quad 0]^T, \boldsymbol{w}_2 = [0 \quad 1 \quad 0]^{\mathrm{T}}, \boldsymbol{w}_3 = [0 \quad 0 \quad 1]^T$
 6:         $\boldsymbol{w} \leftarrow \max_{\boldsymbol{w}_i} \boldsymbol{w}_i^T \Sigma \boldsymbol{w}_i$ for $i$ = 1,2,3
 7:     else
 8:         Compute the three vectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, $\boldsymbol{w}_3$
 9:         if any vector component in $\boldsymbol{w}_i < 0$ for $i = 1, 2, 3$ then
10:             $\emptyset \leftarrow \boldsymbol{w}_i$
11:             $\boldsymbol{w} \leftarrow \max_{\boldsymbol{w}_i} \boldsymbol{w}_i^T \Sigma \boldsymbol{w}_i$ for $i$ = 1,2
12:     Use $\boldsymbol{w}$ in Eq. (3.1) to linear map $\boldsymbol{M}$ to a single band greyscale $\boldsymbol{Q}$
13:     Return $\boldsymbol{w}$, $\boldsymbol{Q}$
```

colour corrected using the greyworld assumption) whereupon using Eq. (3.38) would result in three vectors at infinity. To understand why this happens, we need to look back at Eq. (3.7). It is trivial to show that when $\boldsymbol{\mu}_R = \boldsymbol{\mu}_G = \boldsymbol{\mu}_B$ Eq. (3.7) reduces to Eq. (3.8) and this means that our planes now overlap. When this occurs we say that they have become *degenerate*. This degeneracy increases the dimensionality of our solution space from a line to planar region described by Eq. (3.8). If our optimisation includes the cubic boundary conditions then our solution space is a 2-simplex and our solution can be found at one of its vertices (as we described in Section 3.2.1). This can be shown by solving for the intersection of the half-spaces defined by Eq. (3.4) in the basis of E. The vertices of Eq. (3.8) are

$$\boldsymbol{w}_1 = [1 \quad 0 \quad 0]^T, \quad \boldsymbol{w}_2 = [0 \quad 1 \quad 0]^T, \quad \boldsymbol{w}_3 = [0 \quad 0 \quad 1]^T, \tag{3.39}$$

which mean that a single image band (red, green or blue) is a candidate for our greyscale (line 5 in Algorithm. 2). As before, we use Eq. (3.5) to determine which vector produces the largest variance for our maximum contrast brightness preserving greyscale.

3.4 Brightness Preserving Colour-to-Greyscale: Enhanced Dynamic Range

In the previous section, we described a closed form solution to the brightness preserving greyscale of Qiu et al. (2008). This solution was subject to unit-cubic boundary conditions which with Eq. (3.25) ensured no pixel clipping in the greyscale. The shortcomings of this optimisation occur for images that do not possess full dynamic range. While still outputting a maximum variance solution subject to the cubic constraints, we were producing a greyscale image with a variance lower than that of one which was allowed to possess full dynamic range.

To increase the dynamic range of our images we need to allow our band weights to take on values greater than 1 and less than 0, which we do by removing the cubic boundary conditions. To ensure that we get no image clipping in our greyscale we therefore need to impose the following inequality on each pixel

$$0 \leq \alpha \boldsymbol{R}_i + \beta \boldsymbol{G}_i + \gamma \boldsymbol{B}_i \leq 1. \tag{3.40}$$

Each pixel 'i' in the colour image will therefore have two inequalities associated with it and this ensures that our linear mapping outputs a greyscale that has pixels values between 0 and 1. In the absence of our previous inequalities that described the unit cube, our band weights can now take on values less than 0 and greater than 1. These inequalities represent half-spaces ($\boldsymbol{w}^T\boldsymbol{h}$) whose intersection we will now describe. We start by defining our half-spaces

$$P_i = \{\boldsymbol{w} \in \mathbb{R}^3 : \boldsymbol{w}^T\boldsymbol{h}_i = k \quad | \quad i = 1, 2, ..., 2n, \quad 0 \leq k \leq 1\}, \tag{3.41}$$

which subject to our inequality constraints defines a new set of half-spaces

$$\Pi_i = L \cap P_i. \tag{3.42}$$

As described in Section 3.2.2, the half-spaces that contribute to their intersection satisfy

$$\bigcup \Pi_i = \mathrm{conv}(\Pi_i^+), \tag{3.43}$$

where the superscript (+) defines the dual points of the half-spaces.

Let us look at the simple case, where we have three half-spaces defined by the planes that make up the unit cube faces $\alpha = \beta = \gamma = 0$. By defining our half-spaces of the form $\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \leq 0$ (Franco P. Preparata, 1985), we can map our half-spaces in $\mathbb{R}^3$ to L

$$\boldsymbol{e}_i^T \boldsymbol{n} x + \boldsymbol{e}_i^T \boldsymbol{i} - b_i \leq 0, \tag{3.44}$$

which becomes

$$-\,\boldsymbol{n}(i)\ x - \boldsymbol{i}(i) \leq 0 \qquad i = 1, 2, 3, \tag{3.45}$$

where a single vector component $i \in \boldsymbol{n}$ defines one of the three half-spaces in L with non-homogeneous dual points $(3\boldsymbol{n})$. Our two candidate points are found by taking the convex hull of the dual of the points mapped to L, which in a single dimension becomes a sorting operation

$$\boldsymbol{d} = \left\{ \begin{array}{l} d_{\max} = \max\left\{3\boldsymbol{n}(i) : i = 1, 2, 3\right\} \\ d_{\min} = \min\left\{3\boldsymbol{n}(i) : i = 1, 2, 3\right\} \end{array} \right\}. \tag{3.46}$$

Having filtered out the redundant half-space, we can solve for x in Eq. (3.45). Mapping x back to $\mathbb{R}^3$ and adding on our translation vector $\boldsymbol{i}$ leads to

$$\boldsymbol{w}_i^T = -\frac{1}{\boldsymbol{d}(i)} \boldsymbol{n}^T + \boldsymbol{i}^T \quad i = 1, 2 \tag{3.47}$$

and this produces two band weighting vectors $\boldsymbol{w}$. One of these vectors will maximise the quadratic form (Eq. (3.5)) which is the solution we seek to the brightness preserving colour-to-greyscale problem.

Let us now return to the case where every pixel in the image represents two half-spaces. As before we map them to L (see Algorithm. 3 line 11)

$$\boldsymbol{h}_i^T \boldsymbol{n} x + \boldsymbol{h}_i^T \boldsymbol{i} - b_i \leq 0 \qquad i = 1, 2, ..., 2n \tag{3.48}$$

and this results in $2n$ dual points (where n is the number of pixels in the image)

$$\boldsymbol{d} = \left\{ \begin{array}{c} d_{max} = \max \left\{ \frac{\boldsymbol{h}_i^T \boldsymbol{n}}{\boldsymbol{h}_i^T \boldsymbol{i} - b_i} : i = 1, 2, ..., 2n \right\} \\ d_{min} = \min \left\{ \frac{\boldsymbol{h}_i^T \boldsymbol{n}}{\boldsymbol{h}_i^T \boldsymbol{i} - b_i} : i = 1, 2..., 2n \right\} \end{array} \right\}. \tag{3.49}$$

Again by substituting $\boldsymbol{d}$ into Eq. (3.47) gives us our two band weighting vectors in $\mathbb{R}^3$. As before, we choose the vector that maximises Eq. (3.5) as our band weights to produce our desired greyscale $\boldsymbol{J}$.

The pseudocode for the enhanced dynamic range brightness preserving colour-to-grey is shown in Algorithm. 3 and begins on line 8 for the case we have just described. Lines 8 and 9 defines and maps our half-spaces to the basis of L. Line 10 computes the vertices of the intersection of half-spaces defined by the union of our inequalities. As these half-spaces form points on our line L, the convex hull of our dual-points are found by finding the maximum and minimum points (Eq. (3.49)), and this step is carried out in Algorithm. 1 (line 5). The vertices are then calculated using Eq. (3.47) (line 8 of Algorithm. 1) and mapped back to our original vector space (line 11 of Algorithm. 3).

Lines 5 to 8 addresses the degenerate case which we shall discuss in the next subsection.

Algorithm 3 Enhanced Dynamic Range Brightness Preserving Colour-to-Grey.

```
 1: procedure enhancedBrightnessPreservingColour2Grey
 2:     Input: $\boldsymbol{M}$ as an array of pixel inequalities P, $\boldsymbol{\Sigma}$
 3:     Output: Greyscale image $\boldsymbol{J}$
 4:     $\boldsymbol{i}^T \leftarrow [1/3, 1/3, 1/3]$                         ▷ Interior point
 5:     if $\mu_R = \mu_G = \mu_B$ then
 6:         $\boldsymbol{\Gamma} \leftarrow \boldsymbol{N}$                      ▷ Eq. (3.50)
 7:     else
 8:         $\boldsymbol{\Gamma} \leftarrow \boldsymbol{n}$                      ▷ Eq. (3.33)
 9:     $\Pi_i \leftarrow [P_i\boldsymbol{\Gamma}, (P_i\boldsymbol{i} + b_i)]$   ▷ $\Pi = P \cap L$
10:     $\boldsymbol{v}_i \leftarrow$ intersectHalfspaces($\Pi$)          ▷ Algorithm. 1
11:     $\boldsymbol{w}_i^T \leftarrow \boldsymbol{\Gamma}^T\boldsymbol{v}_i + \boldsymbol{i}^T$   ▷ $L \rightarrow \mathbb{R}^3$
12:     $\boldsymbol{w} \leftarrow \max_{\boldsymbol{w}_i} \boldsymbol{w}_i^T \Sigma \boldsymbol{w}_i$
13:     Use $\boldsymbol{w}$ in Eq. (3.1) to linear map $\boldsymbol{M}$ to a single band greyscale $\boldsymbol{J}$
14:     Return $\boldsymbol{w}$, $\boldsymbol{J}$
```

3.4.1 Degenerate Case

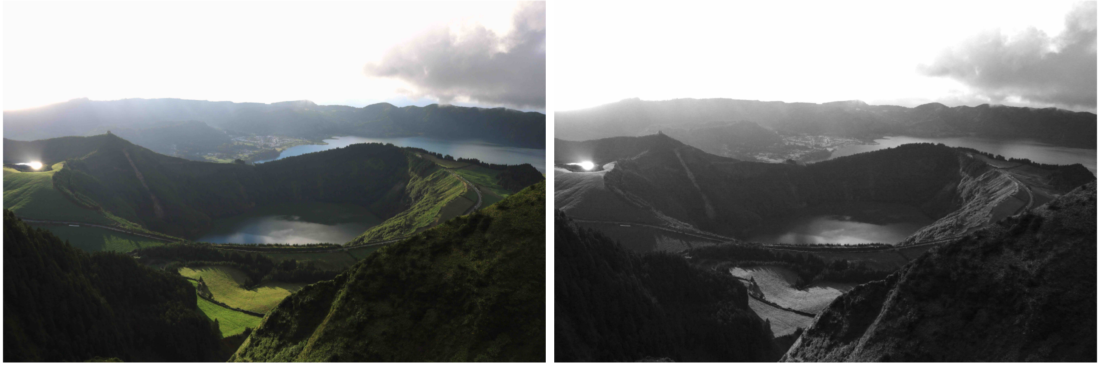
As for our closed form solution that uses the cubic boundary conditions, we must address the situation when our equality constraints are degenerate; a situation that occurs when $\boldsymbol{\mu}_R = \boldsymbol{\mu}_G = \boldsymbol{\mu}_B$. As we have previously described, we have removed the cubic constraints on our band weights and instead posed inequality constraints on each pixel. We must now solve for the intersection of half-spaces (Eq. (3.41)) defined by these inequality constraints in the basis of E (see Algorithm. 3 line 6).

Using the tools of linear algebra, an orthogonal basis in E that spans the null space of our unitary vector $\boldsymbol{u}$ can be computed to be

$$\boldsymbol{N} = \mathrm{Null}(\boldsymbol{u}) = \begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ 0 & -\frac{\sqrt{6}}{3} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}. \tag{3.50}$$

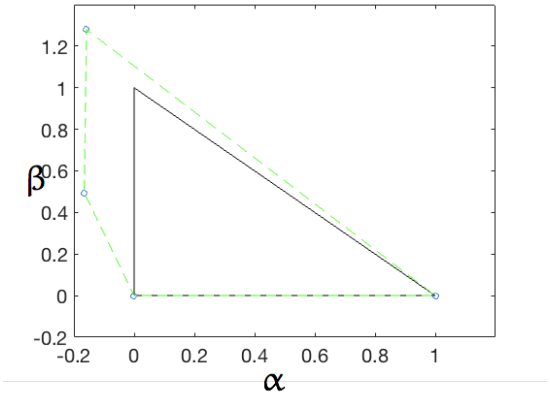
By swapping out $\boldsymbol{n}$ with $\boldsymbol{N}$ in Eq. (3.44) we can produce a mapping of our half-spaces in the basis of E. As our half-spaces are now described by 2 dimensions, we can no longer use sorting to discriminate against our redundant half-spaces and instead we rely on a convex hull algorithm. The convex hull labels which half-spaces form edges of our convex set in addition to their ordering with respect to a circle. We solve for the intersection for their boundary intersections by computing line intersections between two edges of the convex set, before moving onto the next edge. After all the vertices have been computed, we perform the inverse mapping to bring them back to $\mathbb{R}^3$ to obtain a library of coefficient vectors $\boldsymbol{w}$. As before, selecting the vector that gives the highest variance from Eq. (3.5) is the one we use for our brightness preserving greyscale.

Figure. 3.2 shows an example image whose bands all possess the same statistical mean of 0.309 (without gamma correction). In this case the solution to the optimisation of Qiu et al. (2008) would be the vertex of E (subject to the cubic boundary conditions) with the highest variance, which for Figure. 3.2(a) is the red band (see Figure. 3.2(b)).



(a) Colour image (b) Red band

(c) Green band (d) Blue band



(e) Convex set caused by the (green) intersection of half-spaces for each pixel and the (black) cubic constraints.

Figure 3.2: (a) Raw colour photo (`https://www.wesaturate.com/photo/9HG77ydMR2`) that has been colour corrected using the Grey World assumption. The mean of for all bands is 0.309. (b) red band with a variance of 0.402 and range of 1.000, (c) green band with a variance of 0.386 and a range of 0.968 and (d) the blue band with a variance of 0.390 and a range of 0.966.

Figure. 3.2(e) shows the geometry (looking down the γ axis) of the optimisation constructed by Qiu et al. (2008) (black triangle) and our enhanced dynamic range optimisation (green dashed convex set). As the dynamic range for all image bands is not equal to one, the convex set that is produced from our intersection of half-spaces can produce vertices that give band weights greater than 1 and less than zero. For this particular image, Qiu et al.'s (2008) solution is the band with the largest variance; which for the image in Figure. 3.2(a) is the red band (Figure. 3.2(b)). For our enhanced dynamic range optimisation, solely using the red band is also the solution we seek ($\alpha = 1, \beta = 0$ in Figure. 3.2(e)), as it has a larger variance than the greyscales produced using the band weights defined by the vertices for the convex set produced by the intersection of half-spaces. If each band of the image had a range equal to one, then the convex set defined by the intersection of half-spaces for our enhanced dynamic range solution would have reduced to the black triangle and our solution could be computed using our closed form solution to Qiu et al.'s optimisation.

3.5 Results and Discussion

Figures. 3.3, 3.4, 3.6 and 3.7 show greyscale examples $\boldsymbol{Q}$ and $\boldsymbol{J}$. Image (d) in these figures illustrate the areas of contrast change between the two optimisations which has been computed by the ratio

$$\frac{\boldsymbol{J}}{\boldsymbol{Q}+\boldsymbol{J}}, \tag{3.51}$$

where greyscale pixels lower and higher than that of one at 0.5 (see the scale to the right side of the contrast difference image) respectively show darker and brighter regions in $\boldsymbol{J}$ over that of $\boldsymbol{Q}$.

The subfigures (a) and (b) of Figures. 3.5 and 3.8 show the geometry (in band weighting space) of our optimisation where we are looking down the $\boldsymbol{\gamma}$ axis. The black triangle represents set E subject to the cubic boundary conditions. The green dotted convex polytope represents the intersection of half-spaces in the basis of E for our enhanced dynamic range solution. The red dotted line segment represents L, the intersection of our equality constraints.

These figures show the increase of magnitude of the line segment, over that of the optimisation of Qiu et al. (2008) which reaches the edge of the black triangle only. This increase in line segment magnitude can allows for new line segment end points that gives us RGB band weights for a greyscale with enhanced dynamic range and contrast (see Figures. 3.3, 3.4, 3.6 and 3.7). Subfigures (c)-(f) of Figures. 3.5 and 3.8 shows the greyscale histograms between the method of Qiu et al. (2008) and our enhanced dynamic range solution. As we can see, by increasing the dynamic range of the optimisation of Qiu et al. (2008) we see a smoothing and of the distribution of the bin population. The variance, entropy and band weights are quantitatively shown in Table. 3.1 accurate to three decimal places. As can be seen we have increased not just the variance (K) but also the image entropy (S). To understand why this is, we need to consider the entropy of a multivariate Gaussian (Cover and Thomas, 2006)

$$S = \frac{1}{2}\log(2\pi e)^{d}\left|\mathbf{\Sigma}\right|, \tag{3.52}$$

where $|\mathbf{\Sigma}|$ is the determinant of the $d\times d$ covariance matrix and e is the exponential. For a single greyscale the covariance is simply the variance K of the image. So with respect to Eq. (3.52), as our $\boldsymbol{J}$ mapping has a larger variance than $\boldsymbol{Q}$ we are also increasing its entropy and information content.

Quantitatively our $\boldsymbol{J}$ mapping shows an approximate increase of variance and entropy for ‘The Scream’ and ‘Sunrise’ of 20% and 2% respectively. ‘Pool Balls’ exhibits a much lower increase in variance and entropy; approximately 5% and 0.4% respectively. This result can be explained by looking at Figure. 3.8(a) where we see our enhanced dynamic range line end points barely surpassing the black triangle boundary (whose intersection defines the band weights of Qui’s solution) which results in a small increase in magnitude of the line segment. ‘Monarch’ gives a 2% increase in entropy and a 38% increase in variance. Qiu’s solution for this image

Table 3.1: Brightness preserving colour-2-greyscale band weights, variance (K) and entropy (S).

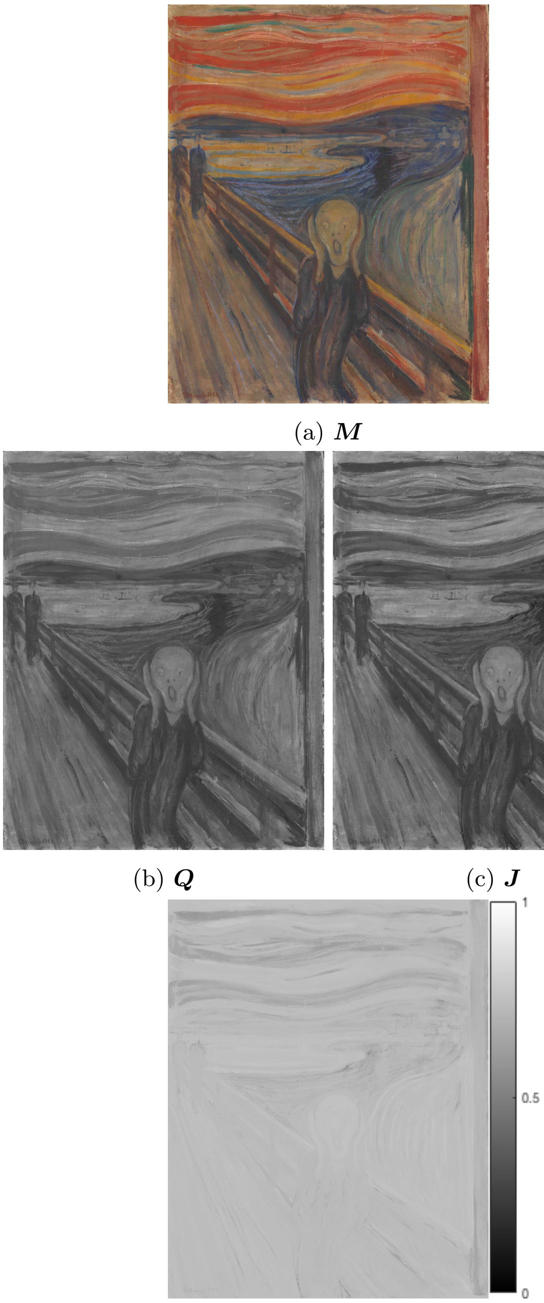
Image	Qiu et al. (2008), $\boldsymbol{Q}$					Enhanced Dynamic Range, $\boldsymbol{J}$				
	α	β	γ	K	S	α	β	γ	K	S
The Scream	0.199	0.801	0.000	0.009	6.552	0.009	1.466	-0.475	0.011	6.696
Sunrise	0.857	0.143	0.000	0.015	7.005	0.989	0.095	-0.084	0.0182	7.142
Pool Balls	0.437	0.563	0.000	0.038	7.262	0.446	0.584	-0.031	0.040	7.290
Monarch	0.258	0.742	0.000	0.026	7.208	0.514	-0.645	1.131	0.036	7.345

lies at the top most line segment end point in Figure. 3.8(b). For our enhanced dynamic range solution the optimum vertex switches to the lowest line segment end point. Although this results in a large increase in image variance, it also introduces contrast inversion on the wing, where the yellow shades in the colour image become very dark grey values in our enhanced greyscale (see Figure. 3.7(c)). The result from Qiu's optimisation remains a more accurate representation of the colour image, where these pixels on the wings are assigned brighter values. The remaining images are consistent between the two optimisations with respect to greyscale ordering.

3.6 Complexity Analysis

Our solutions to the brightness preserving colour-to-greyscale all rely on arithmetic calculations in addition to rational number sorting or convex hull functions. Arithmetic calculations have a linear time complexity whereas number sorting and convex hull functions vary depending on which algorithm is used. We implement all of our algorithms in the technical computing language MATLAB. For the convex hull and sort functions, MATLAB uses the versatile divide and conquer recursive algorithms: 'quickhull' and 'quicksort' that have time complexities of $O(n^2)$ (Franco P. Preparata, 1985, O'Rourke, 1998). The rational number sorter 'mergesort' is well suited for our purposes and has a time complexity of $O(n \log n)$, lower than that of quicksort. The quickhull is a generic and well accepted algorithm which can be applied to problems with dimensionalities greater than two. For our application we will only ever need to compute the convex hull in the plane (when our band means are degenerate) which allows us to use either of the $O(n \log n)$ time complex algorithms 'monotone chain' and 'Graham scan' algorithms. As these algorithms possess time complexities greater than that of the linear arithmetic calculations they become the limiting operations for our solution.

Qiu et al. (2008) use the quadratic programming function in MATLAB which for the interior-point method has a time complexity of $O(n^3)$ (YE and TSE, 1989). This time complexity is greater than the solution we present which if we use the mergesort and monotone chain results in an overall time complexity of $O(n \log n)$.



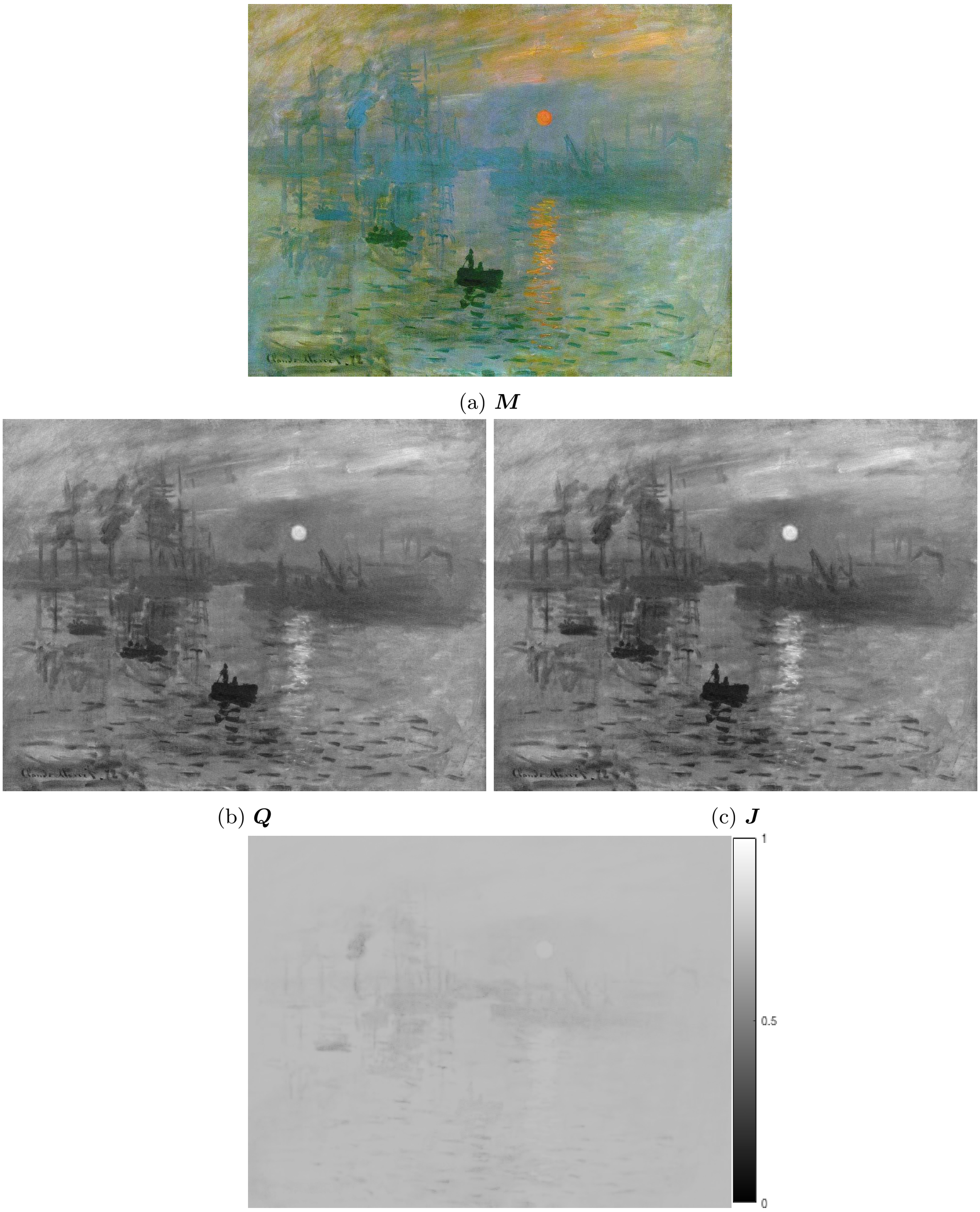
(a) $\boldsymbol{M}$

(b) $\boldsymbol{Q}$

(c) $\boldsymbol{J}$

(d) $\boldsymbol{J}/(\boldsymbol{Q}+\boldsymbol{J})$

Figure 3.3: (a) The Scream and (b) the brightness preserving colour-to-grey of Qiu et al. (2008), (c) our enhanced dynamic range brightness preserving colour-to-grey and (d) the differences in contrast between the two mappings.



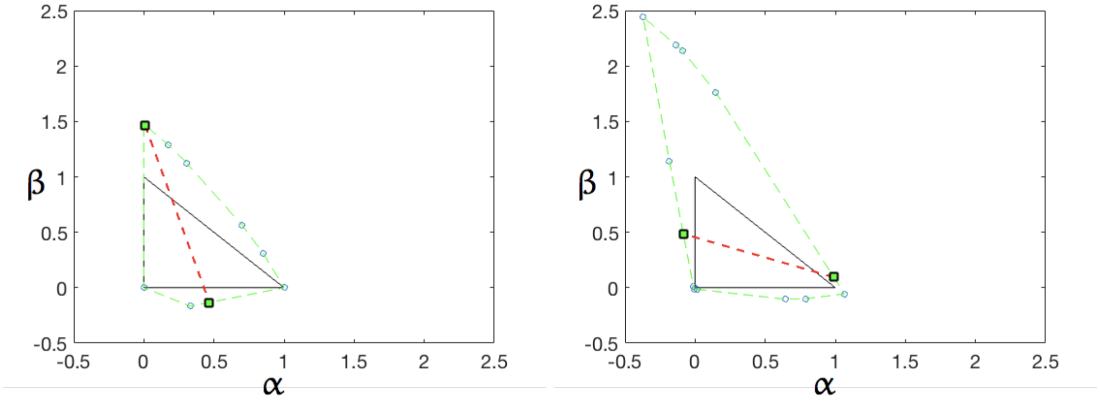
(a) $\boldsymbol{M}$

(b) $\boldsymbol{Q}$

(c) $\boldsymbol{J}$

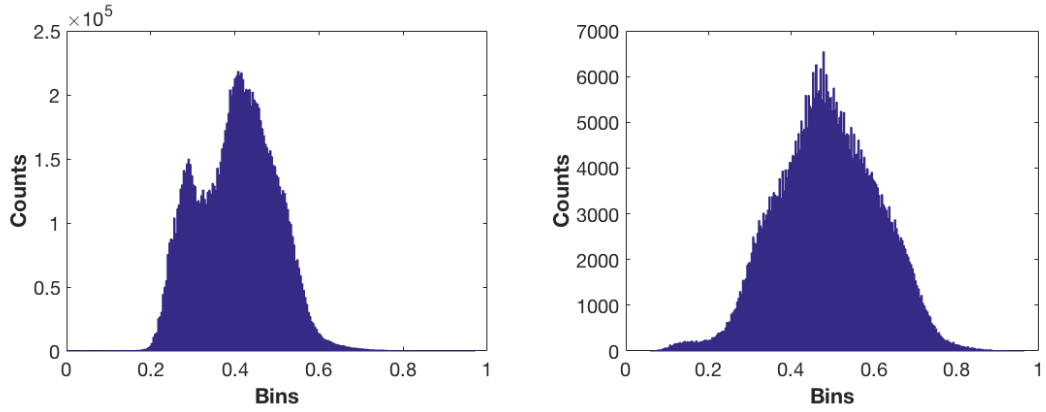
(d) $\boldsymbol{J}/(\boldsymbol{Q}+\boldsymbol{J})$

Figure 3.4: (a) Monet's sunrise and (b) the brightness preserving colour-to-grey of Qiu et al. (2008), (c) our enhanced dynamic range brightness preserving colour-to-grey and (d) the differences in contrast between the two mappings.



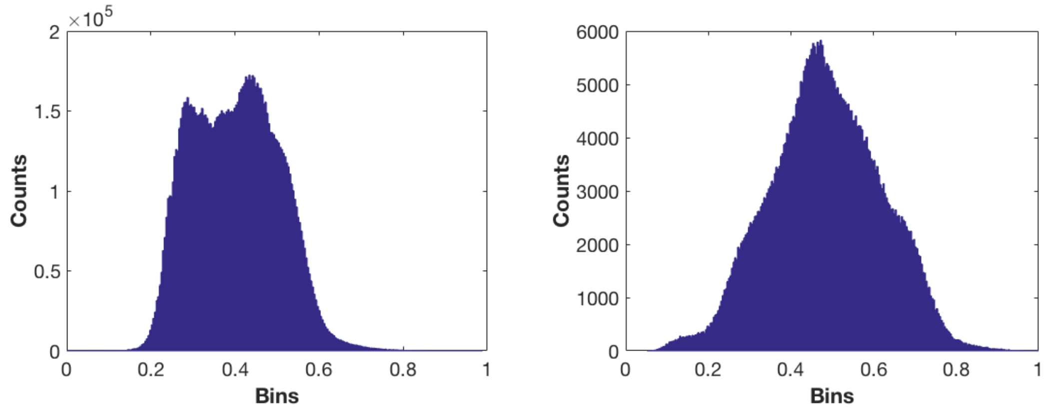
(a) The Scream: coefficient space

(b) Sunrise: coefficient space



(c) The Scream histogram: $\boldsymbol{Q}$

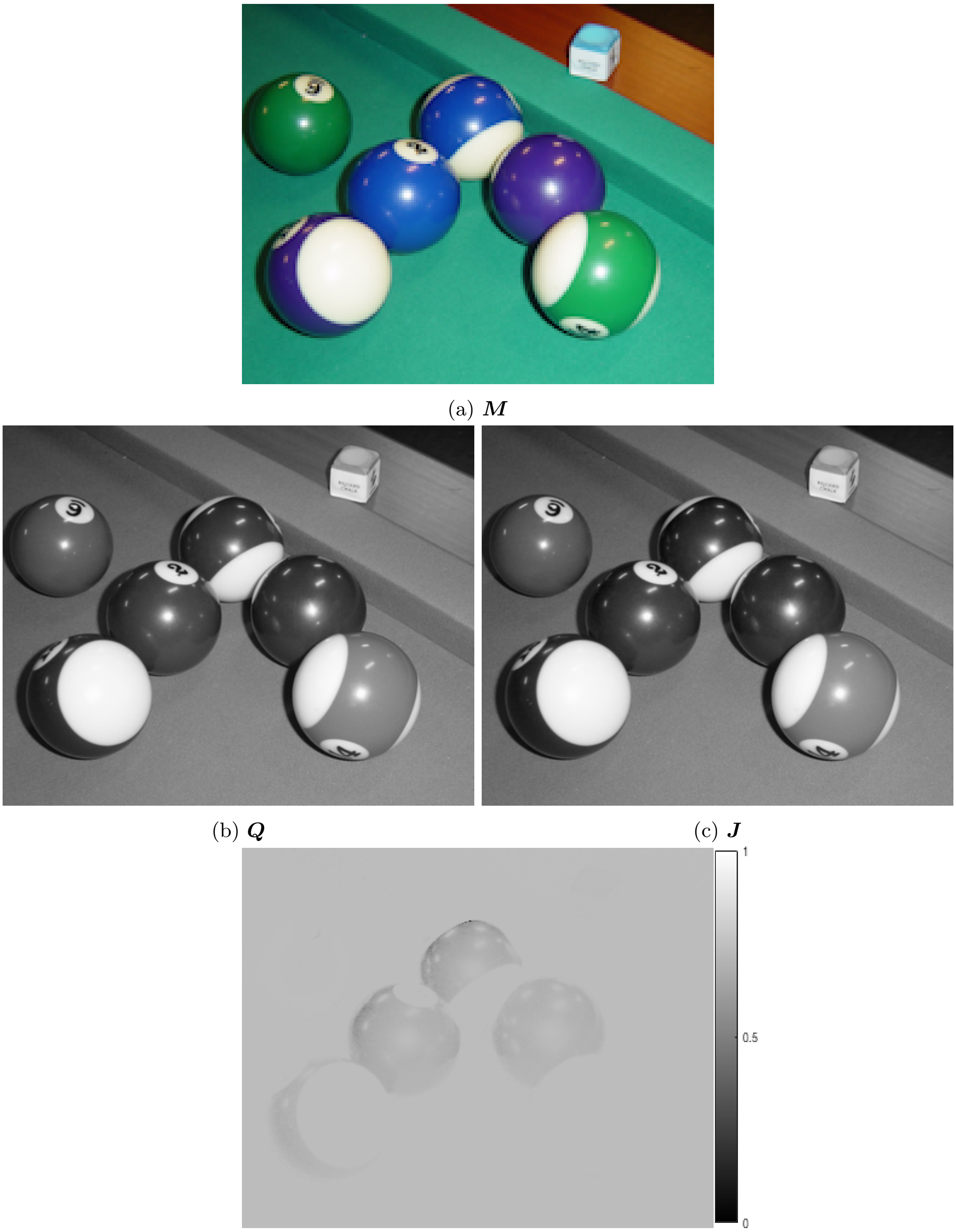
(d) Sunrise histogram: $\boldsymbol{Q}$



(e) The Scream histogram: $\boldsymbol{J}$

(f) Sunrise histogram: $\boldsymbol{J}$

Figure 3.5: Looking down the γ axis of coefficient space for the images (a) The Scream and (b) Sunrise. (c) The image histograms from Qiu et al. (2008) for The Scream and (d) Sunrise. The image histograms from our enhanced dynamic range optimisation for (e) The Scream and (f) Sunrise.



(a) $\boldsymbol{M}$

(b) $\boldsymbol{Q}$

(c) $\boldsymbol{J}$

(d) $\boldsymbol{J}/(\boldsymbol{Q}+\boldsymbol{J})$

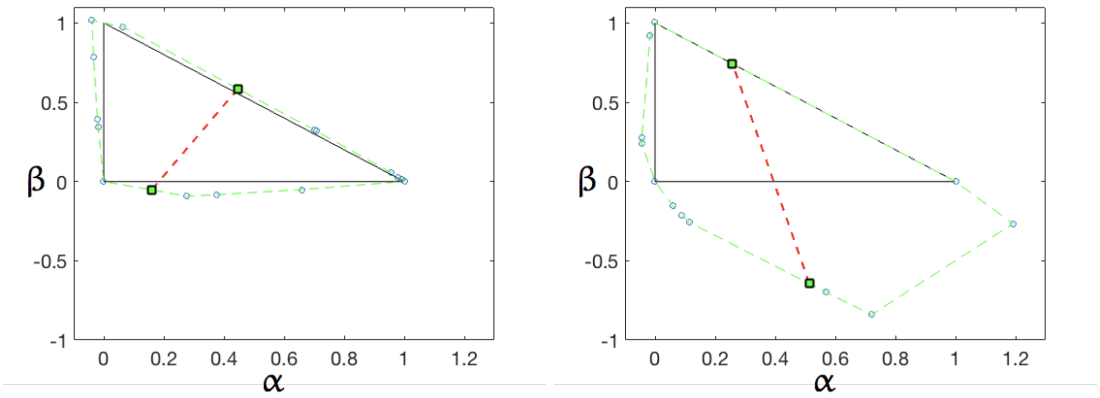
Figure 3.6: (a) PoolBalls colour image taken from Cadik's dataset, (b) its brightness preserving colour-to-grey of Qiu et al. (2008), (c) its enhanced dynamic range brightness preserving colour-to-grey and (d) the differences in contrast between the two mappings.

(a) $\boldsymbol{M}$

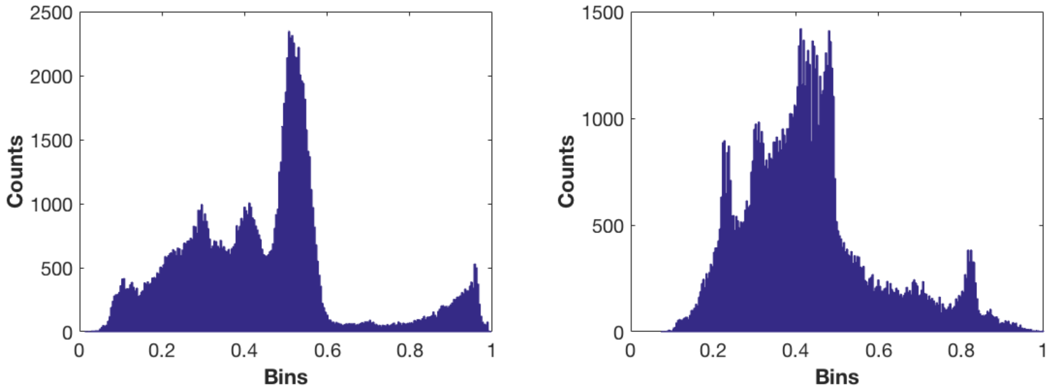
(b) $\boldsymbol{Q}$ (c) $\boldsymbol{J}$

(d) $\boldsymbol{J}/(\boldsymbol{Q}+\boldsymbol{J})$

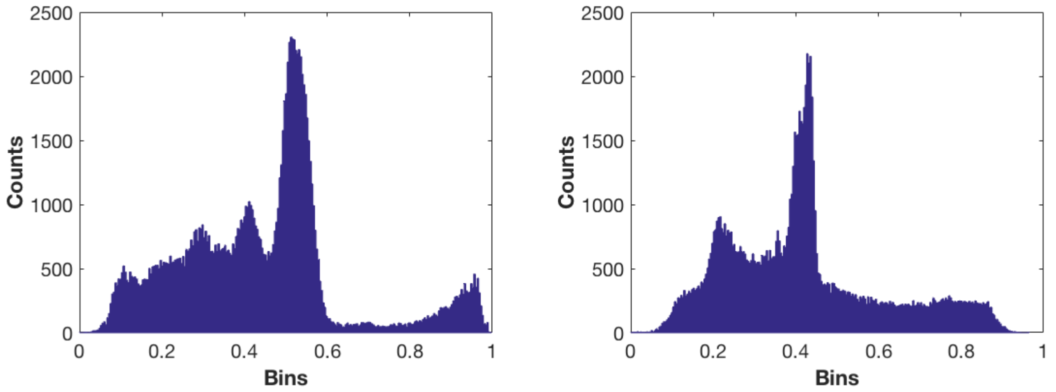
Figure 3.7: (a) Monarch taken from Cadik's dataset, (b) its brightness preserving colour-to-grey of Qiu et al. (2008), (c) its enhanced dynamic range brightness preserving colour-to-grey and (d) the differences in contrast between the two mappings.



(a) PoolBalls: coefficient space (b) Monarch: coefficient space



(c) PoolBalls histogram: ***Q*** (d) Monarch histogram: ***Q***



(e) PoolBalls histogram: ***J*** (f) Monarch histogram: ***J***

Figure 3.8: Looking down the γ axis of coefficient space for the images (a) PoolBalls and (b) Monarch. (c) The image histograms from Qiu et al. (2008) for PoolBalls and (d) Monarch. The image histograms from our enhanced dynamic range optimisation for (e) PoolBalls and (f) Monarch.

3.7 Conclusion

We started this chapter by showing that a vertex of a convex set is the solution to maximising positive-definite quadratic forms. Consequently, this formed the foundation of our closed form solution to the brightness preserving colour-to-greyscale problem. For the optimisation of Qiu et al. (2008), this convex set is typically a line segment whose vertices can be found using the band means in standard arithmetic calculations or by solving for the intersection of three half-spaces. The vertex of the convex set that provides the greatest image variance is found by a sorting operation, the computation that bounds the time complexity of the solution. When using the mergesort algorithm, the complexity of our solution is $O(3\ \log(3))$.

Using the optimisation of Qiu et al. (2008) on colour images that do not possess full dynamic range results in greyscale images that also lack full dynamic range. To maximise the dynamic range of our greyscales we remove the cubic boundary conditions and instead add two inequality constraints to every pixel to ensure no image clipping. This optimisation relies on solving for the intersection of half-spaces to solve for our new line segment. Again, the line segment end points are the candidates to the solution to our enhanced dynamic range brightness preserving colour-to-greyscale with the final solution found by rational number sorting, which if the mergesort algorithm is used results in a time complexity of $O(n\ \log n)$, where n is the number of inequalities used in the optimisation (twice the number of pixels).

Both solutions presented in this chapter have lower time complexities than using the quadratic programming solver in MATLAB ($O(n^3)$). In the next chapter we shall bench-mark our greyscale images produced from our enhanced brightness preserving greyscale solution by psychophysical preference experiment.

Chapter 4

Colour-to-Greyscale: Evaluation

In the previous chapter we described a closed form solution to the brightness preserving greyscale problem. Subject to unit cube boundary conditions, this solution was found by arithmetic calculation followed by sorting three rational numbers. By removing the cubic constraints on the band weights, we expanded on the number sorting solution by constraining each individual pixel to the dynamic range of the display. This allowed for greyscales with greater contrast for colour images that were deficient in their dynamic range.

In this chapter we seek to ascertain whether our brightness preserving greyscale method produces visually pleasing greyscales. To carry out this task we employ psychophysics on results collected from a participant based preference experiment: a tournament of image comparisons between different colour-to-grey methods.

The outline of this chapter follows a standard experiment based research methodology. We begin by introducing the topic of psychophysics in the context of comparative judgement. We then progress to the sections: method, results and discussion and lastly a conclusion.

4.1 Psychophysics

In the field of colour-to-grey, it has become common practice (Connah et al., 2007, Montagna, 2011, Harris, 2015) to rank images from the scores obtained in a preference experiment. From this competitive tournament, colour-to-greyscale algorithms can be ranked by their ability to produce visually pleasing greyscales.

There is an abundance of algorithms that convert colour images into their respective greyscale. Depending on the desired criteria, some algorithms perform better than others for certain image types and vice versa. At the fundamental level, the criteria that a viewer uses when deciding on how well an algorithm performed when producing a greyscale depends on their own unique psychology and it is for this reason that we evaluate our technique using psychophysics.

A typical means of applying psychophysics to acquire a metric of preference is to apply Thurstone's law of comparative judgement. Thurstone believed that when two stimuli are perceived each one produces psychological response (S_A and S_B) in the human mind. When both stimuli are presented together multiple times in a discriminatory manner, the psychological response can be modelled by a normal distribution whose mean value $S_A - S_B$ provides a measure of which stimuli "won", see Figure. 4.1. Here we see an example where the discriminate difference is equal to 1, which tells us S_A produce a greater psychological stimulation than S_B over multiple observations. Either side of the 0 represents the probability that $P(S_A > S_B)$ (right) or $P(S_B > S_A)$ (left) was preferred. Under certain assumptions it is possible to generate a scale of which stimuli produces the greater response without the observer ever having to specifically quantify the discriminate difference but by simply making a choice of preference.

There are numerous cases for which Thurstone's law of comparative judgement can be applied (Engeldrum, 2000 and Gescheider, 1997). We use Case V as it is typically used in the imaging literature especially with respect to colour-to-greyscale preference experiments. The major underlying assumption for this case is that the standard deviation for the distributions of each stimuli are equal. When assigned an arbitrary value of 1 results in the distribution $S_A - S_B$ having a standard deviation of $\sqrt{2}$. We shall now show how we apply Case V to our work.

Throughout the course of our experiment, we are measuring the proportion of time that $P(S_A > S_B)$ or $P(S_B > S_A)$. By measuring the frequency of "wins" an algorithm collects throughout our tournament of comparisons allows us to apply Gaussian statistics to compute this scale of proportionality. These measured values should follow the normal cumulative distribution

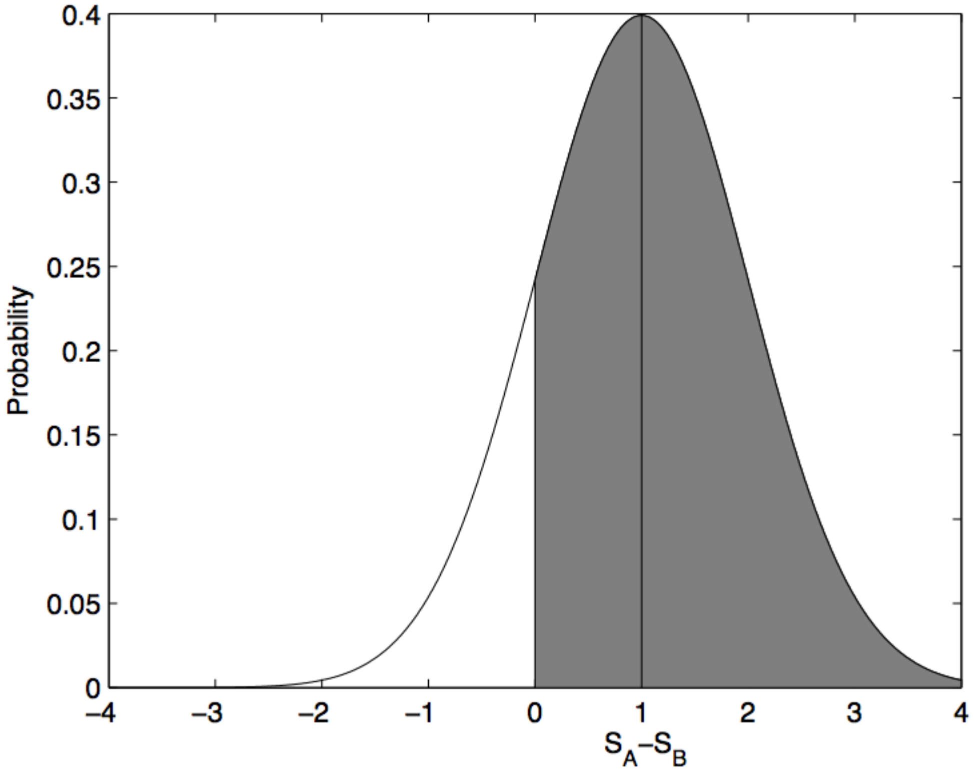


Figure 4.1: The shaded area represents the probability $P(S_A > S_B)$ that the response to stimulus S_A is greater than that to a stimulus S_B (Montagna, 2011).

function H. That is to say, we assume

$$\begin{aligned} H(S_A - S_B) =& P(S_A > S_B) \\ =& \frac{1}{2\sqrt{\pi}} \int_0^{+\infty} \exp\left(-\frac{1}{2}\left(\frac{t-(S_A-S_B)}{\sqrt{2}}\right)^2\right) dt, \end{aligned} \tag{4.1}$$

where $S_A - S_B$ is the mean and $\sqrt{2}$ the standard deviation (Case V assumption) of the distribution and t is the number of algorithms. It is now possible to calculate the scale value differences $S_A - S_B$ from P by inverting $H(\cdot)$

$$S_A - S_B = H^{-1}(H(S_A - S_B)) = H^{-1}(P(S_A > S_B)). \tag{4.2}$$

For the sake of simplicity we have only described the scenario of comparing two different stimuli. To extend the model to our experiment where we are comparing four, we tabulate the data in a proportion matrix

$$P = \begin{bmatrix} P(S_1 > S_1) & \cdots & P(S_1 > S_4) \\ \vdots & \ddots & \vdots \\ P(S_4 > S_1) & \cdots & P(S_4 > S_4) \end{bmatrix}, \tag{4.3}$$

which if we substitute into Eq. (4.2) results in the score matrix gives us

$$S = H^{-1}(P) = \begin{bmatrix} S_1 - S_1 & \cdots & S_1 - S_4 \\ \vdots & \ddots & \vdots \\ S_4 - S_1 & \cdots & S_4 - S_4 \end{bmatrix}. \tag{4.4}$$

From the score values we can compute the final score for each row j by summing along the columns i of our proportion matrix

$$\frac{1}{t}\sum_{i=1}^{t}(S_j - S_i) = S_j - \bar{S} \qquad j = 1, 2, ..., t. \tag{4.5}$$

Making the assumption that the mean scale value $\bar{S} = 0$ then allows us to calculate the scale value for each treatment S_j.

4.2 Preference Experimental Method

In this section we shall describe the method where we gauge the effectiveness of our enhanced brightness preserving colour-to-greyscale against competing pre-existing colour-to-grey algorithms.

To conduct this experiment we first needed to gather greyscale imagery from the algorithms we wished to benchmark ours against. We did this by emailing the authors of the published work reviewed in Chapter 2 for 22 greyscales produced from their algorithms using what has now become the standard dataset in the colour-to-grey literature: `http://cadik.posvete.cz/color_to_gray_evaluation/`. From those that responded, we selected three algorithms that are based on linear mappings from the colour cube model: GRU (Grundland and Dodgson, 2005), ALS (Alsam, 2009) and SNG (Song et al., 2013). We imposed this selection process as we wished to benchmark our method against similar competing techniques that potentially could have application to panchromatic image sharpening. We decided against using the greyscales outputted from the optimisation of Qiu et al. (2008) on the basis that our enhanced method would either replicate or produce superior (with respect to contrast and information content) greyscales depending on whether their RGBs were full dynamic range or not.

32 anonymous volunteers were recruited via word of mouth to participate in our image preference experiments (one participant per experiment). Each volunteer was verbally checked for colour normal vision before being allowed to participate. No other discriminatory measures were used in the participant selection process.

The experiment was conducted in a dark room run by the Colour Laboratory at the University of East Anglia. To avoid eye strain and unwanted reflections, the room was dimly lit by overhead lamps in addition to a globe lamp (6500K) that was situated behind a calibrated (sRGB colour space) HP LP2480zx PC monitor. Participants were allowed to freely adjust the viewing angle and distance to the monitor based on their own comfort.

The image preference experiment comprises of comparisons of two greyscale images by an observer. For a single comparison two greyscale images of the same scene were shown that had been generated by 2 out of the 4 methods: GRU, ALS, SNG and our own enhanced dynamic range brightness preserving colour-to-greyscale, UEA. Participants were asked to consider both images, and to select one using the computer the experiment was being ran on. The explicit instructions will be: "Which greyscale image do you prefer?" No further direct instructions were given, rather they were left free to choose their own criteria on which to form their judgements. The results collected over the course of 32 experiments were analysed using standard psychophysical statistical analysis. At the end of an experiment, each participant was given the opportunity to provide feedback about how they went about defining their preference or about the experiment itself.

For the sake of participant attention span, the experiment was broken down into two sessions: 11 different images per session with a break of at least three hours between them. The greyscale comparisons were made between images of the same scene only. As there are 4 methods of doing this greyscale conversion, this results in $N(N-1)/2$ comparisons for a single image, where N is the number of' "versions" of the same image to be tested (in this case $N = 4$). Each pair of

Table 4.1: Frequency matrix collected over the entire dataset for all participants.

Algorithm	ALS	GRU	SNG	UEA	Total
ALS	0	560	555	641	1756
GRU	848	0	683	713	2244
SNG	853	725	0	712	2290
UEA	767	695	696	0	2158

versions will be compared twice over the course of the experiment, one iteration of version A on the left, B on the right, and one of A on the right and B on the left. This means that a single participant will make 264 comparisons over the course of both sessions (a single experiment).

4.3 Results and Discussion

Throughout the course of the experiment, participant choices are stored in a frequency matrix. Table. 4.1 shows the frequency matrix collected over all 32 participants for all 22 images. The rows of this table catalogue the total wins during a comparison with the columns. For example, over the entire experiment, GRU beat ALS 848 times. The total number of wins are stored in the final column where our algorithm (UEA) came third place with 2158 wins out of a total of 8448 comparisons.

Thurstone's Case V was applied to a frequency matrix comprising of all data entries collected during our psychophysical experiment to produce the final score values for each algorithm, see Figure. 4.2. The error bars were calculated from the normal distribution at the 95% upper (+) and lower (−) confidence interval

$$x \pm 1.96 \times \frac{\sqrt{2}}{\sqrt{n}}, \tag{4.6}$$

where n is the number of observations (4224 for each algorithm over all images and participants). Due to the significant overlap of the error-bars between GRU and SNG we conclude that there is no clear decisive winner for producing greyscales for this dataset. We can say that our algorithm came third and ALS was the least preferred overall. Participant feedback would generally report poorly on the increased texture the ALS method would produce, often referring to its 'muddiness' and 'artefacts'.

Figures. 4.3 - 4.6 shows the algorithm score values for each image, with 95% confidence

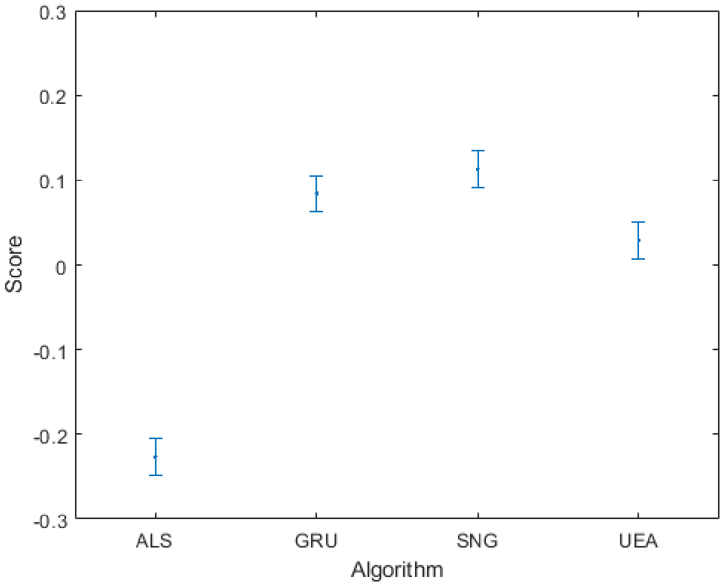


Figure 4.2: Score values for each algorithm over the entire experiment.

interval error bars. As is expected, there is dispersion among the favoured algorithm across the image dataset. Our maximum contrast brightness preserving algorithm performs particularly well for the synthetic images, IM2, C8TZ7768 and ramp. It lost every comparison for the synthetic 'ColorsPastel' due to the 'C' having no contrast, making the word unreadable. This particular image draws attention to a flaw in our enhanced brightness preserving colour-to-greyscale method. The presence of a white border (or even a single pixel) around the image automatically reduces our optimisation to that of Qiu et al. (2008). Now, if a given component i in our weighting vector $\boldsymbol{w}$ approaches one, then the solution will approach a single band from the colour image $\boldsymbol{M}$. The information in the remaining bands is lost in the greyscale mapping. For the chosen image band, pixels that have identical brightness values will exhibit no contrast, or difference, between each other. This problem will result in contrast loss if they are adjacent to one another. We can express this as

$$\lim_{\boldsymbol{w}(i) \mapsto \frac{\boldsymbol{M}_k(\dot{i})}{\boldsymbol{M}_l(\dot{i})} \mapsto 1} \nabla \boldsymbol{J}_{k,l} = 0, \quad i = 1 \text{ or } 2 \text{ or } 3, \tag{4.7}$$

where the subscripts k and l refer to two different RGB pixel triplets. The weighting vector for the image ‘ColorsPastel’ is $\alpha = 0.002, \beta = 0.98, \gamma = 0$; approximately all the information in the mapping is from the green band, which for the ‘C’ and its surrounding contain green band values 0.68 and 0.69 respectively. The information in the red and blue bands is lost, despite them having significantly different values from each other at the two different locations. As a result of this mapping, all contrast is lost for the ‘C’ and its background to an observer.

We also performed poorly for our greyscale of Monet’s Sunrise, where participants favoured a darker, higher contrast night-time setting. This is a point of interest: upon feedback, participant preference was generally made on a compromise between contrast and realism. Participants would favour images that displayed a level of contrast that they felt the original colour image possessed. A number of the participants were ignorant to the painting by Monet and assumed that the colour original depicted a full moon in a nights sky.

Amongst the photos in the dataset (see photos in the Figures. 4.7 to 4.10), UEA scored neutral/positive. It was heavily favoured on the images ‘peppers’ and ‘poolballs’. The only photo it scored negatively on was ‘monarch’ where the butterfly’s wings appeared as a dark-grey. As this occupies the central and focal point of the image, participants felt that there was a general lack of visible detail and contrast and so favoured the treatment of ALS, whose local contrast enhancing approach performed particularly well. GRU and SNG scored consistently around the neutral/positive. The only images where these algorithms scored last were ‘IM2’ for GRU and ‘watch’ for SNG.

Table. 4.2 contains the statistical analysis which serves to provide insight into the behaviour of the participants during the experiment. ‘Agreement’ (u) shows how often participants would choose the same image in a comparison. It ranges from 1 in the case of perfect agreement to -0.042 (Connah et al., 2007). It is possible that purely random selection of preferred images could result in perfectly correlated choices among the participants. To strengthen our confidence in our calculated u, we test the null hypothesis that selections were made at random through the chi squared (χ^2) test. Lastly we looked at participant consistency (ξ), which is a measure of how often a participant would make the same choice when presented with the same comparison. The synthetic images ‘ColorsPastel’ and ’ramp’ scored the highest in agreement ($\sim$ 0.7 and 0.5) and achieved a value greater than 0.9 for consistency. This means there was very little uncertainty about what a participant preferred when confronted with different treatments for

Table 4.2: Summary statistics for each image from the preference experiment. The line in the middle divides session 1 and session 2.

Image	Agreement (u)	chi-square (χ^2)	Confidence	Consistency (ξ)
25ishi	0.2791	111.5	p <0.001	0.8594
peppers	0.2793	111.5625	p <0.001	0.8594
arctichare	0.0152	42.9375	p <0.001	0.8203
poolballs	0.0977	42.9375	p <0.001	0.9453
butterfly	0.0526	25.8750	p <0.001	0.5234
C8TZ7768	0.2320	93.6875	p <0.001	0.8828
ColorsPastel	0.7116	275	p <0.001	0.9609
colorWheel	0.0561	27.1875	p <0.001	0.8828
bouquette	0.0933	41.2500	p <0.001	0.8125
fruit	0.0355	19.4375	p <0.050	0.8594
IM2	0.2774	110.8750	p <0.001	0.9375
poppy	0.1103	47.6875	p <0.001	0.9141
hats	0.0747	34.25	p <0.001	0.8281
monarch	0.1870	76.6875	p <0.001	0.8047
portrait	0.1204	51.5	p <0.001	0.7891
ramp	0.5056	197.1250	p <0.050	0.9688
serrano	0.1123	48.4375	p <0.001	0.8516
ski	0.0073	8.75	0.2> p	0.6875
monet	0.3575	141.1250	p <0.001	0.9063
tree	0.0288	16.8750	p <0.01	0.5469
tulip	0.0645	30.3750	p <0.001	0.8594
watch	0.1237	52.75	p <0.001	0.7734

this image. The more complex photo images showed less agreement, with ‘girl’ scoring a u value of 0.3, however the consistency was high (0.85) which means that the participants were certain in their selection, it just differed between participants based on their own psychology. Participants would often mention the image ‘tree’ as being the hardest to make a clear preference on, as all algorithms would produce a similar looking greyscale. Similarly for the image ‘ski’ , which possesses a multitude of focal points for a participant to use during image comparison. This is clearly evident from Table. 4.2 where they broadly had low coefficient scores. Upon feedback, participants in these instances would claim to often choose randomly or simply change their mind on what image they preferred during the experiment due to a change in their image focal point of reference. This can be seen to be true for the image ‘ski’ where the confidence is above our threshold at the 5% limit which means we can claim the images were chosen at random. A final point of interest from participant feedback concerned the image ‘serrano’. Post experiment,

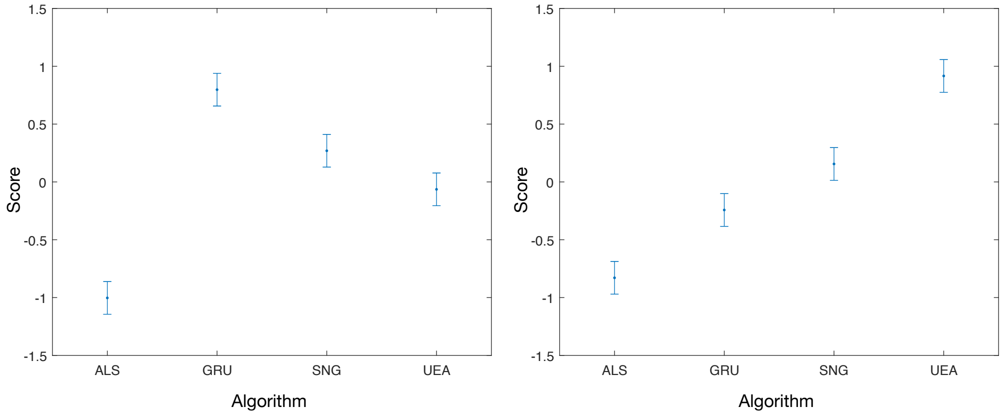
Table 4.3: Tally based on the method ranks order for the dataset.

	1st place	2nd place	3rd place	4th place
ALS	6	1	7	9
GRU	6	6	6	4
SNG	6	8	5	3
UEA	4	7	6	6

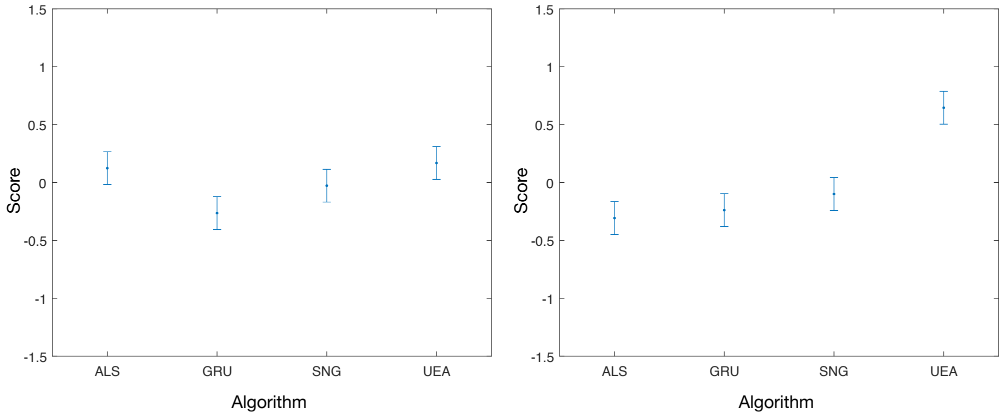
most participants did not know what the picture portrayed until they were presented with the colour original; a point that perfectly illustrates the struggle of greyscale imaging when trying to convey colour information.

Table. 4.3 shows the tally from Figures. (4.3) - (4.6) and provides of how many times an algorithm came first, second, third or fourth. The tally is based solely on the scores and shows the distribution of numbers per place. It can be seen SNG and ALS has the highest tally for the first and last two places, GRU is consistently high from first to third place and that UEA peaks in wins around second and third places.

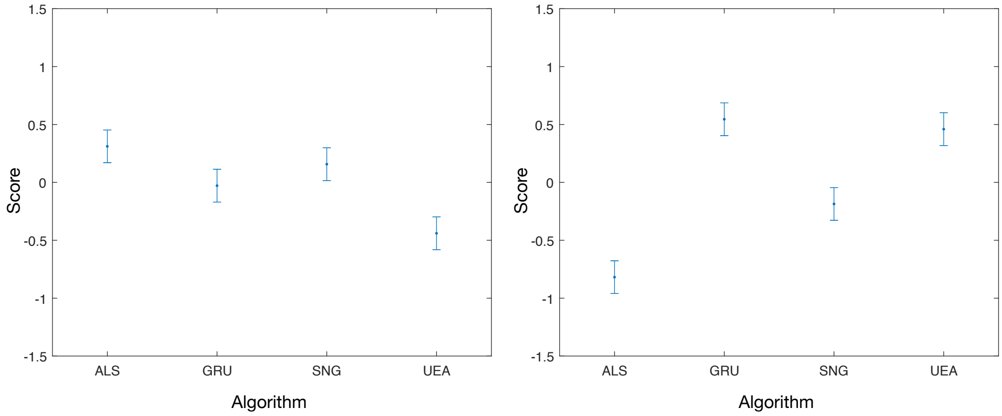
Figures. (4.7) - (4.10) show the original colour and the treated images for each algorithm over both sessions. From visual inspection we can see that the 5 images where our optimisation scored first place (‘peppers’, ‘arctichare’, ‘poolballs’, ‘IM2’ and ‘ramp’) typically possess a low number of hues.



(a) 25ishi (b) peppers

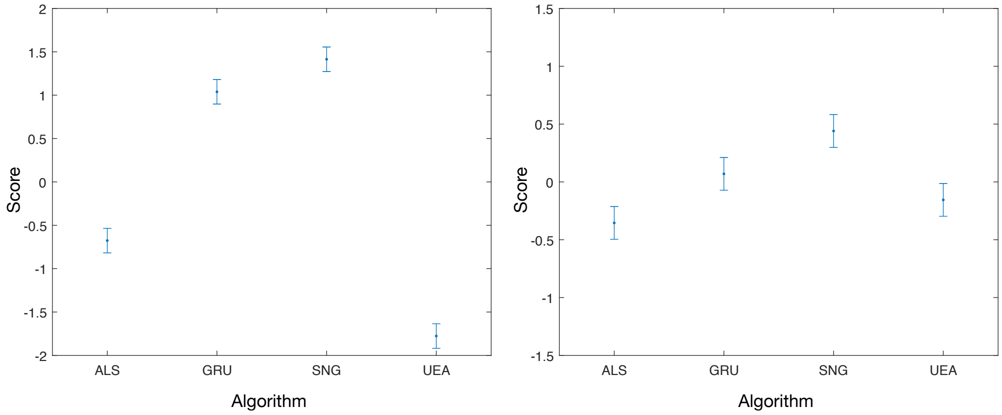


(c) arctichare (d) poolballs

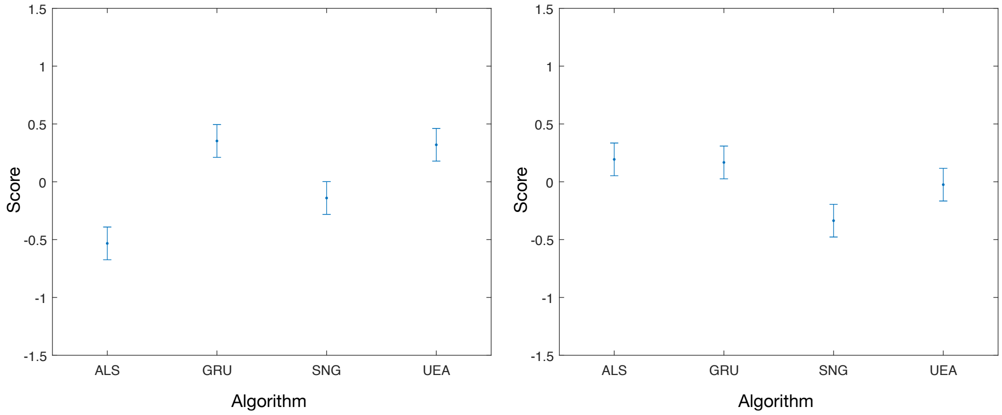


(e) butterfly (f) C8TZ7768

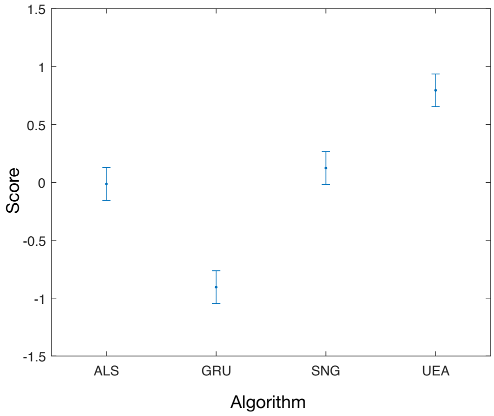
Figure 4.3: Image preference scores over the 32 participants.



(a) ColorsPastel (b) ColorWheelEqLum200

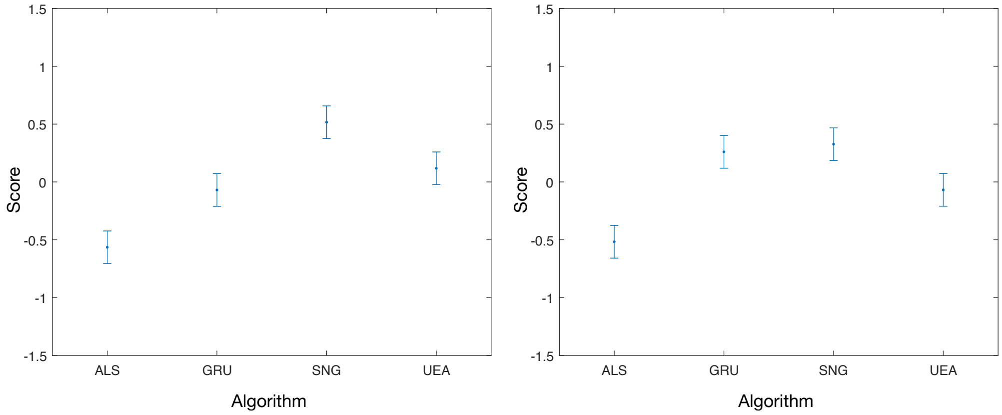


(c) bouquette (d) fruit

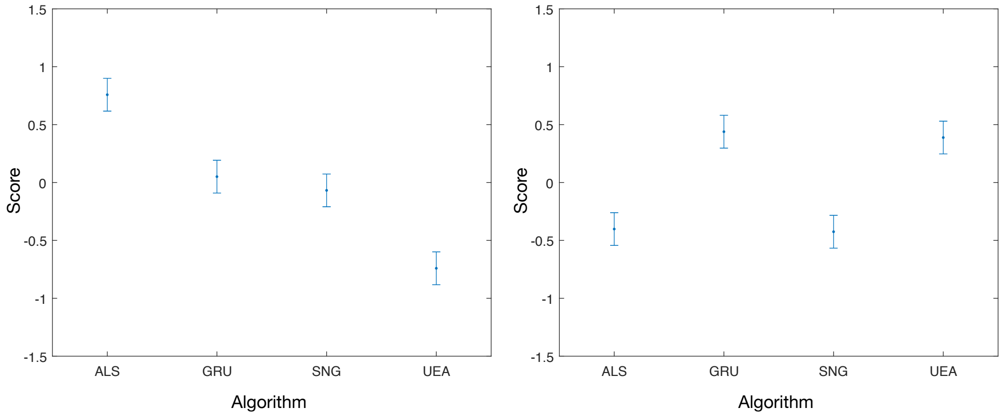


(e) IM2

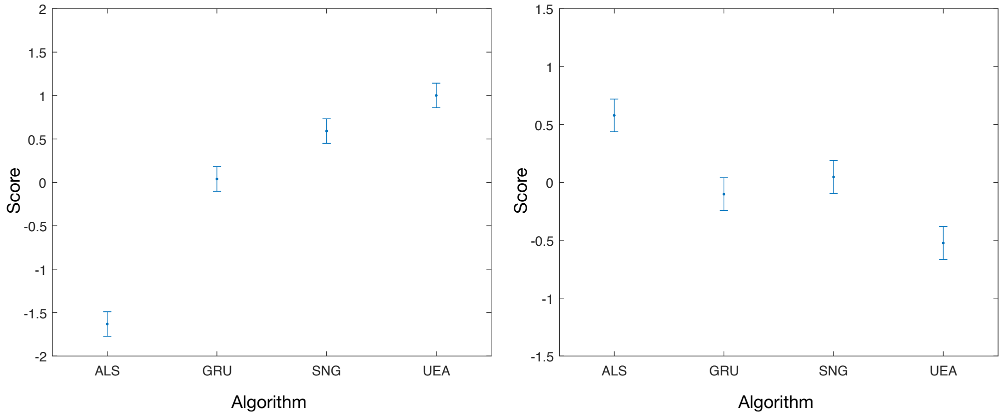
Figure 4.4: Image preference scores over the 32 participants.



(a) poppy (b) hats

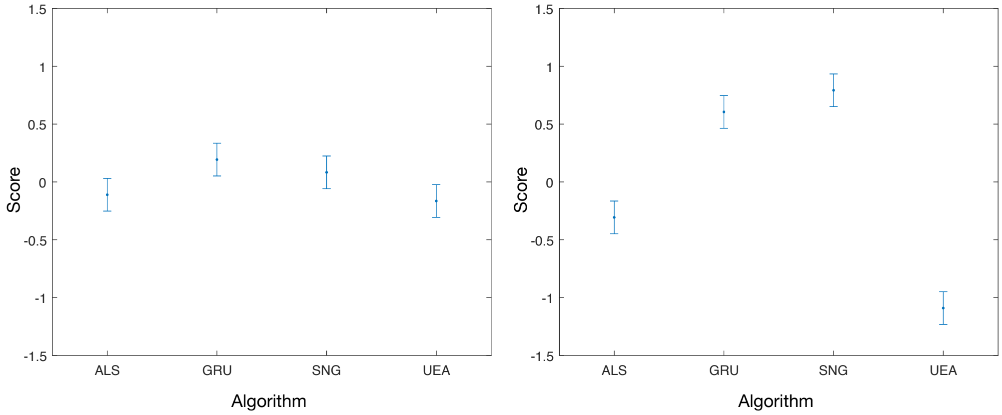


(c) monarch (d) portrait

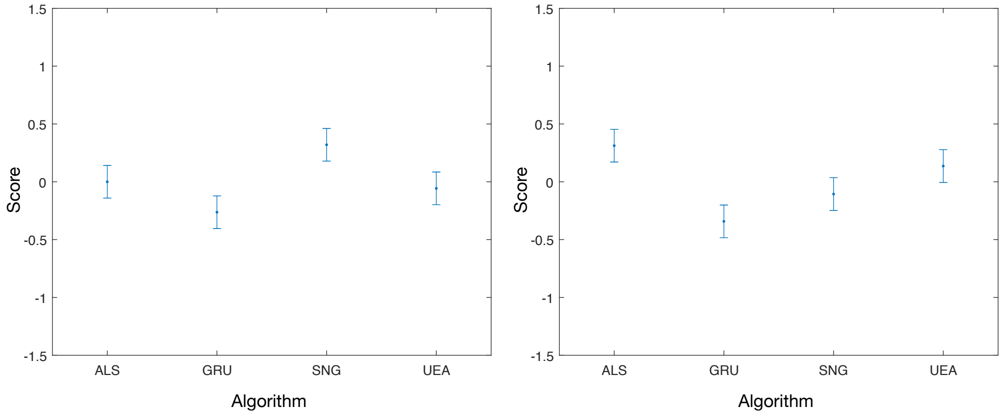


(e) ramp (f) serrano

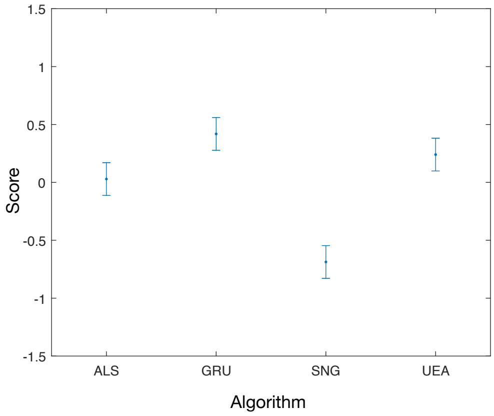
Figure 4.5: Image preference scores over the 32 participants.



(a) ski (b) monet

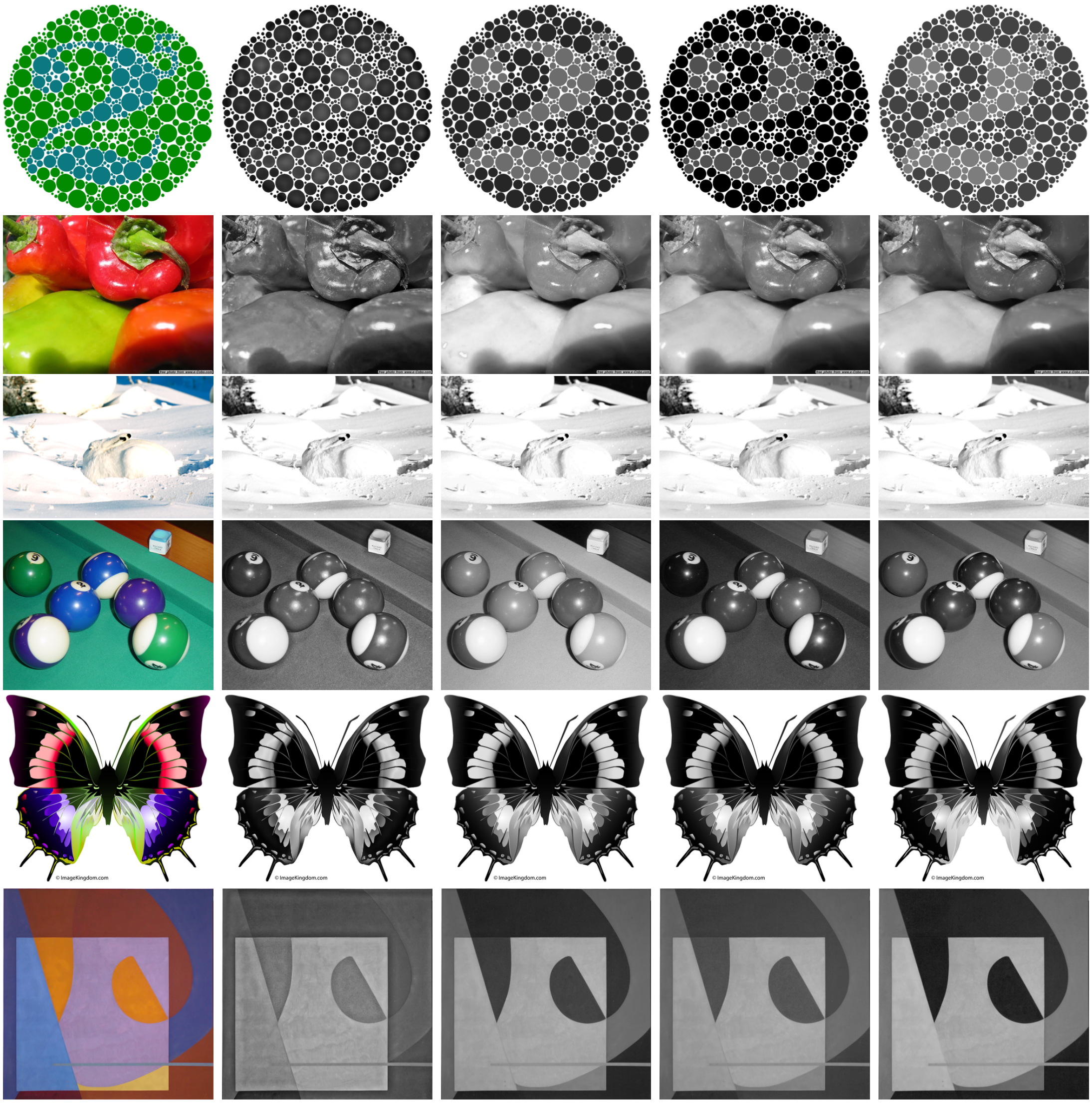


(c) tree (d) tulips



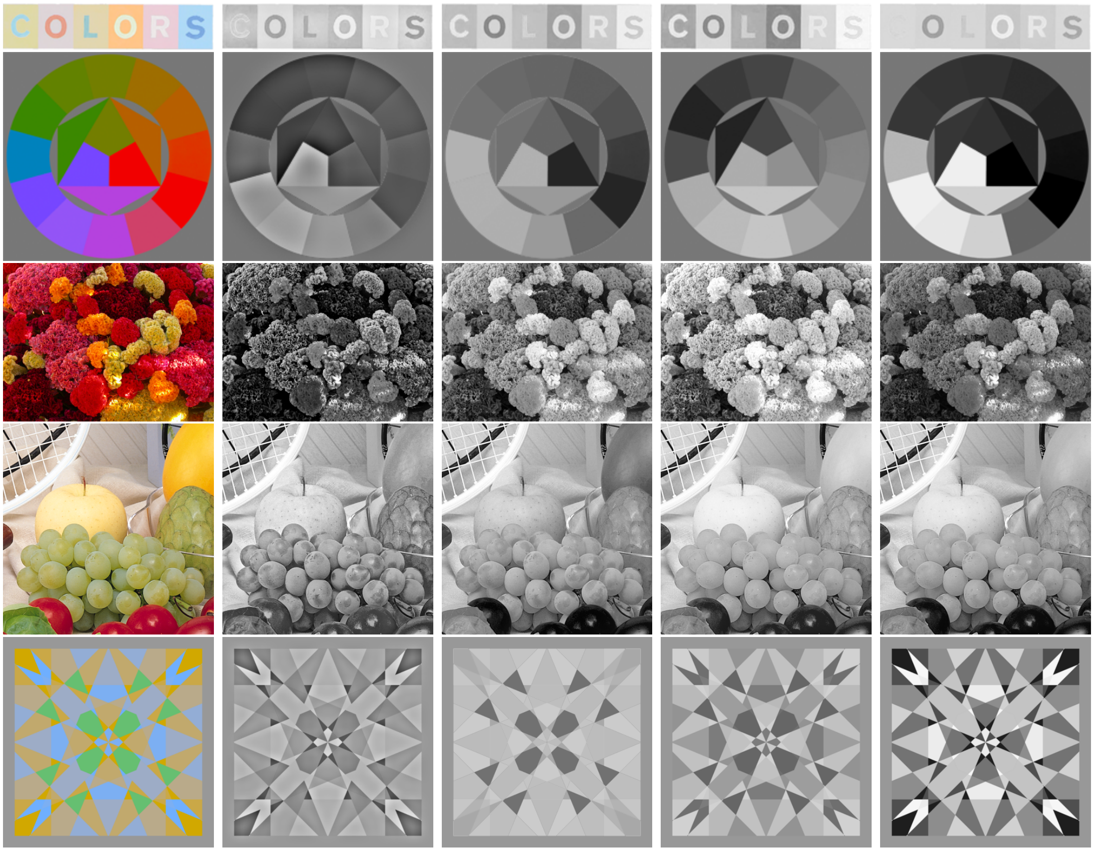
(e) watch

Figure 4.6: Image preference scores over the 32 participants.



RGB ALS GRU SNG UEA

Figure 4.7: Images used in the paired comparison experiment for each algorithm in addition to the original RGB. From top to bottom: 2 ishihara, peppers, arctichare, poolballs, butterfly and C8TZ7768.



RGB ALS GRU SNG UEA

Figure 4.8: Images used in the paired comparison experiment for each algorithm in addition to the original RGB. From top to bottom: ColorsPastel, ColorWheelEqLum200, bouquette, fruit and IM2.

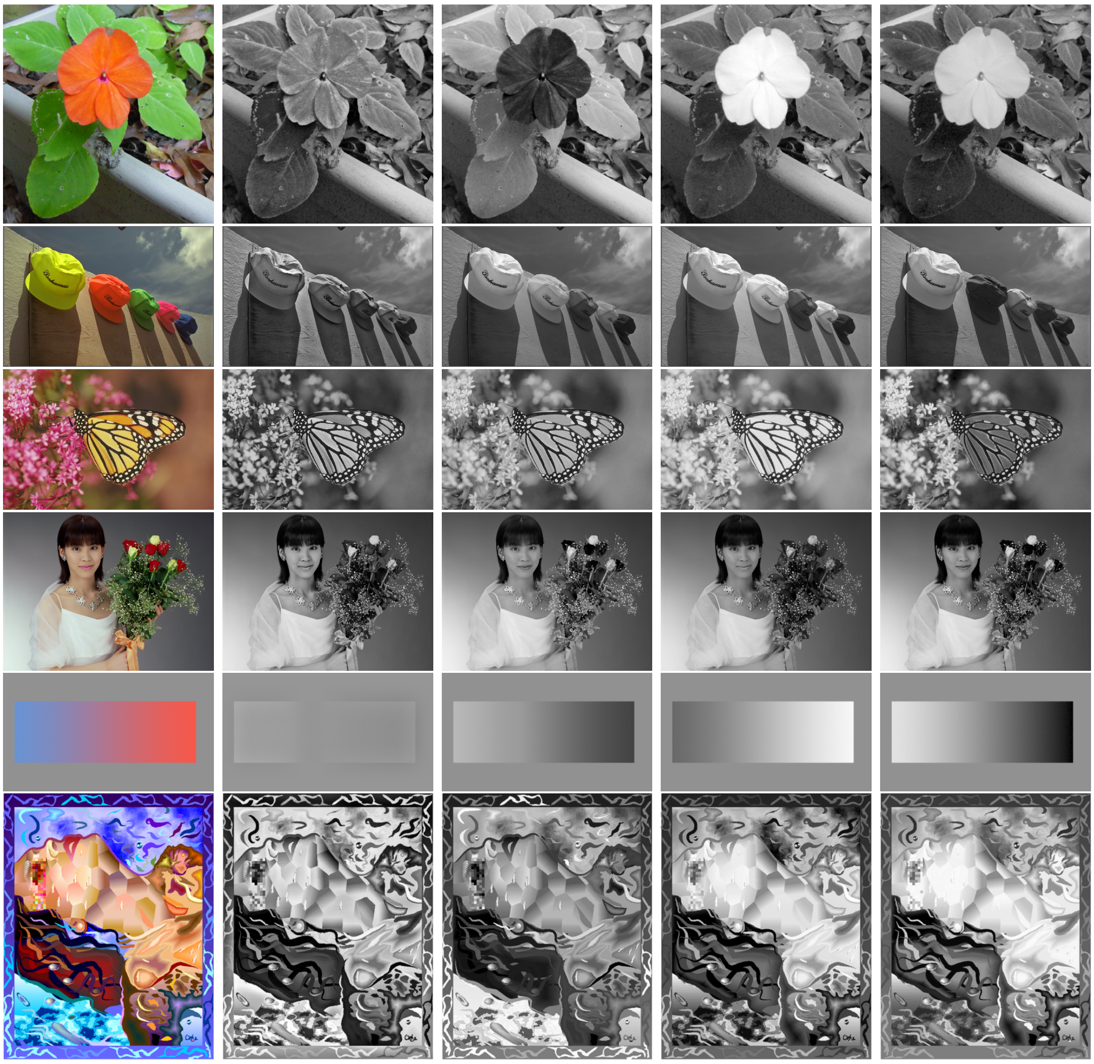


Figure 4.9: Images used in the paired comparison experiment for each algorithm in addition to the original RGB. From top to bottom: poppy, hats, monarch, portrait, ramp and serrano.

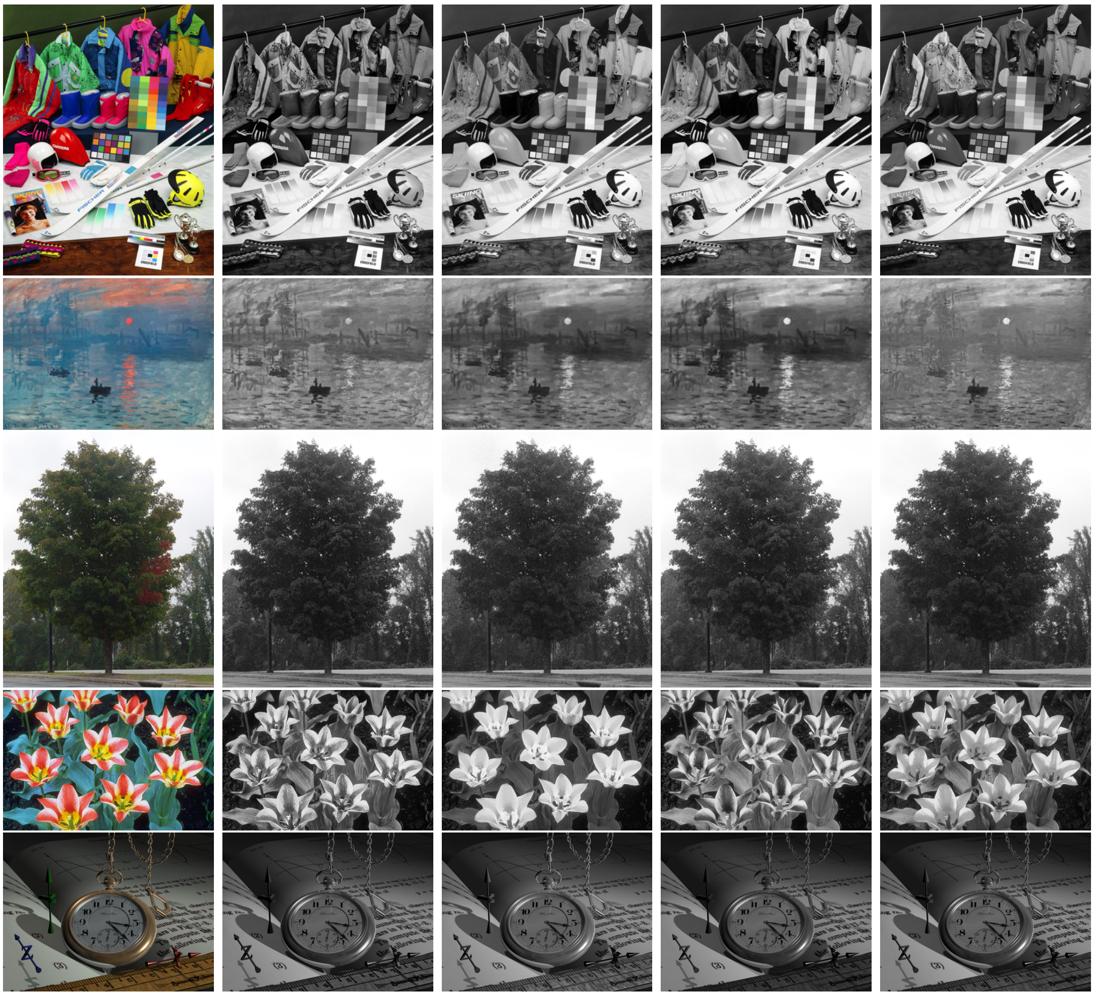


Figure 4.10: Images used in the paired comparison experiment for each algorithm in addition to the original RGB. From top to bottom: ski, monet, tree, tulips and watch.

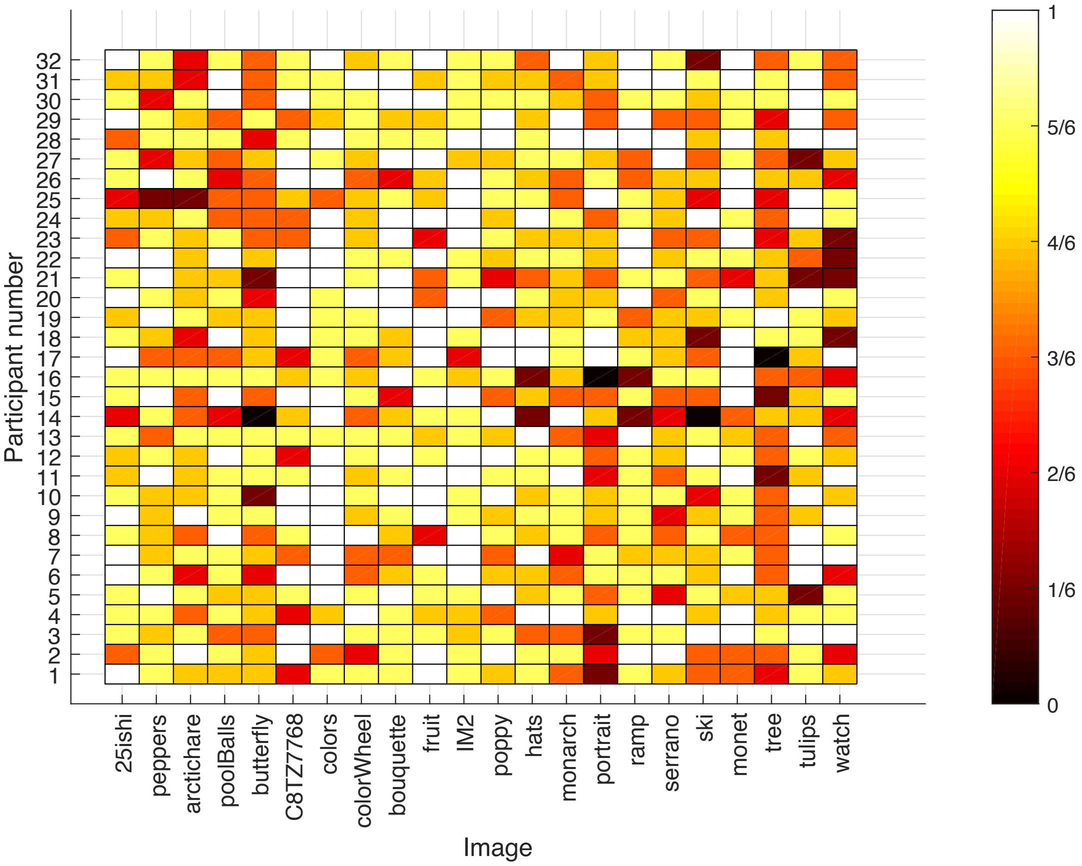


Figure 4.11: Raw consistency for each participant for all images in the dataset. A white square (1 on the key) means that when a participant made am algorithm choice for that particular image, they would always make the same algorithm choice when confronted with the same image combination. A black square (0) would mean they were not consistent at all and they would choose a different algorithm when presented the same image pair.

A detailed depiction of consistency is given in Figure. 4.11 which shows the individual participant consistency per image. White blocks correspond to a value of 1, which means that they were 100% consistent for all algorithm combinations for that image. Conversely, black means that they never chose the same algorithm. The scale from white to black (defined by the key) defines the fractional consistency between 100% and 0%. For example, from this, we can say that for the image ‘tree’ only participant 19 always chose the same algorithms during a comparison for this image.

4.4 Conclusion

In this chapter we described how we evaluated the images produced from four colour-to-grey methods using psychophysical preference experiments. Including our own maximum contrast brightness preserving solution, we chose three other competitive methods based on similar projection based mappings.

The data collected through the participation of 32 volunteers and 22 images in a tournament of comparisons ended with each method receiving a score of preference at the 95% confidence limit. Subject to this confidence interval, our maximum contrast solution was seen to score third out of fourth place. The method whose score came first was based on an optimisation of preserving colour contrast and the lowest came from the local pixel based method, which un-naturally increased texture in many of the greyscale images.

Although many participants chose contrast and image detail to be of paramount importance when making a choice of preference, a portion felt that too much contrast was un-natural and instead preferred a compromise which reflects the contrast that would have been seen in the colour original. Our method tended to score highly on images with a low number of hues and score second place more than other ranks. Despite coming third overall we conclude that an optimisation grounded in maximising contrast is still a viable means of producing visually pleasing greyscales.

This chapter finishes our work in colour-to-greyscale mappings. In the next chapter we shall move onto the topic of panchromatic image sharpening and report on how linear mapping multiband data has application in producing edge maps.

Chapter 5

Panchromatic Sharpening: Literature Review

Having brought an end to our work in colour-to-greyscale mappings, we now seek to apply what we have developed to the field of panchromatic image sharpening. There exists a wealth of journal articles that review existing pan-sharpening techniques. Recently Li et al. (2017) have produced a paper that broadly describes the state of the art of pixel-level image fusion techniques and their applications. Ghassemian (2016) looks at the broader context of image fusion by reviewing techniques used in all three levels; pixel, feature and decision. Vivone et al. (2015) have produced a critical comparison among pan-sharpening algorithms and Amro et al. (2011) published a survey on pan-sharpening with new and classical methods. Lastly it is worth mentioning the work of Pohl and van Genderen (1998) as it is cited frequently in the literature especially when it comes to what are now known as the classical pan-sharpening techniques. Their follow up article Pohl and van Genderen (2014) builds of their previous publications and mostly serves as a source of references for recent advances in the field.

Following the structure of these articles we shall start our review with the classical techniques and move onto how more recent methods attempt to solve their short comings. We shall then move onto frequency domain, multi-resolution and hybrid methods and how they are incorporated in the component substitution framework.

5.1 Component Substitution (CS)

Panchromatic sharpening methods by Component Substitution (CS) are often referred to as the classical techniques and comprise of the Intensity-Hue-Saturation (IHS), Principle Component Analysis (PCA) and Gram Schmidt (GS). Pan-sharpening based on CS is one of the most popular methods to employ due to its simplicity, computational speed and their ability to inject high levels of spatial detail. The technique relies upon swapping out the low resolution spatial information (or part of it) with that of the high resolution panchromatic image $\boldsymbol{P}$. How the low resolution spatial information is acquired from the colour image depends on the transformation done prior to the substitution. The transformation can be viewed as a projection of the image from RGB vector space to a component space that separates out the spatial and spectral information. After the substitution has been made, the inverse transform provides the final pan-sharpened image.

5.1.1 Intensity-Hue-Saturation (IHS)

IHS is the most common of the CS due to its simplicity. The colour image is projected into a space that separates the image into the three components of human colour perception: The intensity component is a brightness map that contains the spatial information in the image, hue relates to the dominant wavelength that contributes to what we see as the colour and saturation provides the measure of the colours purity relative to grey. There is an abundance of ways to decompose an RGB image into IHS space, with the main difference being that of which colour model they are derived from (Smith, 1978). Image intensity calculations based on Smith's Triangle Model is popular within the literature for its accurate undiscriminatory representation of RGB image brightness (Carper et al., 1990, Núñez et al., 1999)

$$\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{V}_1 \\ \boldsymbol{V}_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{2}{6} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{G} \\ \boldsymbol{B} \end{bmatrix}, \tag{5.1}$$

$$\boldsymbol{H} = \tan^{-1}\left(\frac{\boldsymbol{V}_1}{\boldsymbol{V}_2}\right) \tag{5.2}$$

and

$$\boldsymbol{S} = \sqrt{\boldsymbol{V}_1^2 + \boldsymbol{V}_2^2}. \tag{5.3}$$

Substituting out $\boldsymbol{I}$ with the high spatial resolution $\boldsymbol{P}$ followed by an inverse transform back to the original RGB colour space results in the pan-sharpened image. The disadvantage of the IHS technique is that it is limited to just three bands.

To reduce spectral distortions that a full substitution of $\boldsymbol{I}$ with the $\boldsymbol{P}$ can bring, partial replacement of the intensity component can be made to produce a new greyscale $\boldsymbol{J}$. Full replacement simply requires a direct substitution of $\boldsymbol{I}$ with the $\boldsymbol{P}$, whereas a partial replacement is given by (Thomas et al., 2008)

$$\boldsymbol{J} = \rho \boldsymbol{I} + (1 - \rho)\boldsymbol{P}, \tag{5.4}$$

where ρ is the mixing coefficient of the intensity image. When $\rho = 0$ the replacement becomes full and $\boldsymbol{J} = \boldsymbol{P}$.

To overcome having to apply a transform and its inverse, Tu et al. (2001) proposed a Generalised Intensity-Hue-Saturation (GISH) method which simplifies the process to adding injection gains to the RGB vectors. The injection gains can be computed by taking the difference between $\boldsymbol{P}$ and $\boldsymbol{I}$ images. The disadvantage of this technique is that the authors are adding a numerical constant to their RGB vectors which changes their direction resulting in saturation distortion. This method does provide the advantage over standard IHS by allowing for more than three spectral bands to be used in the process by allowing for a simple approximation of the intensity image $\boldsymbol{J}$ by either an equally weighted convex sum (Tu et al., 2001) or from scene specific empirical analysis (Tu et al., 2004) that best represents the degrees of spectral overlap between the multispectral ($\boldsymbol{M}$) and panchromatic bands. The CS method has been summarised (Rahmani et al., 2010, Vivone et al., 2015) with the following injection scheme

$$\boldsymbol{H}_n = \boldsymbol{M}_n + \boldsymbol{g}_n(\boldsymbol{P} - \boldsymbol{J}), \tag{5.5}$$

where $\boldsymbol{J}$ is our greyscale created from a weighted sum of the spectral bands

$$\boldsymbol{J} = \sum_{1}^{d} \omega_n \boldsymbol{M}_n, \tag{5.6}$$

$\boldsymbol{H}_n$ represents the n^{th} high resolution pan-sharpened spectral band, $\boldsymbol{M}_n$ the n^{th} low resolution spectral band captured from the satellite and the vector $\boldsymbol{g}_n$ is a vector that modulates the gains for the n^{th} band.

As we mentioned in Chapter 1, the band weights are typically chosen such that $\boldsymbol{J}$ has the maximum correlation with $\boldsymbol{P}$. Provided there is sufficient correlation, the assumption that $\boldsymbol{J}$ is a low-pass-filtered approximation of $\boldsymbol{P}$ can be made and that by taking their difference $(\boldsymbol{P}-\boldsymbol{J})$ we are extracting an edge map that contains the parts of $\boldsymbol{P}$ not obtainable by convex combination of the spectral bands $\boldsymbol{M}$. In other words, this difference can be regarded as the high resolution spatial details missing from the low resolution multiband image. This process of edge extraction is often used in colour image contrast enhancement and is known as an unsharp mask.

The assumption that $\boldsymbol{P} - \boldsymbol{J}$ produces an unsharp mask for $\boldsymbol{P}$ is only valid to an approximation because it is dependent on the level of spectral overlap and correlation between the panchromatic and multispectral bands. In the instances when the correlation is sufficient then a linear regression is used to determine the band weights ω. Bands that do not have sufficient overlap with the panchromatic can be dismissed; Mascarenhas et al. (1991) excludes both the blue and infrared bands in their regression for their SPOT-1 images and Boggione et al. (2003) and Xu et al. (2008) remove the blue for their Landsat-7 imagery. Using this discriminatory selection process, Xu et al. (2008) fused multispectral bands taken from six additional satellites. From these additional six, only two had the range in the union of their multispectral bands that had sufficient overlap with the panchromatic: QuickBird and IKONOS. To further minimise the least squares error in their $\boldsymbol{J}$, Xu et al. (2008) allowed for an additional constant to the convex sum of $\boldsymbol{M}$.

Instead of calculating the band weights for the upsampled $\boldsymbol{M}$ bands that best approximate $\boldsymbol{P}$, Aiazzi et al. (2007) proposed applying a low-pass filter to $\boldsymbol{P}$ and then calculating ω that would approximate this downsampled panchromatic image. These weights were then applied to the upsampled $\boldsymbol{M}$ bands under the assumption that the coefficients would be nearly identical had the spectral bands been captured at the full resolution as that of $\boldsymbol{P}$. They refer to this as adaptive GISH/IHS or GISHA/IHSA.

Garzelli et al. (2008) elegantly solve for band dependent gains by minimising the mean square error (MMSE) between each spectral band and the panchromatic. This produces an edge map for each band. The MMSE optimisation was solved using the minimum variance unbiased estimator.

Rahmani et al. (2010) compute their band weights using a constrained linear regression. Using the Lagrangian Multiplier method, they constrain their weights to be non-negative. To further insure the quality of their unsharp mask, they use an edge-detecting function to mitigate distortion of the low-frequency components of $\boldsymbol{M}$. As the band weights for $\boldsymbol{J}$ and its edge map are image dependent they refer to the method as Adaptive-IHS (AIHS).

To avoid drastically changing the low-frequency components of $\boldsymbol{M}$, Choi et al. (2011) applies partial-replacement (Eq. (5.4)) to produce high resolution spectral bands. Appreciating that the quality of fusion relies heavily on the correlation of $\boldsymbol{M}$ with $\boldsymbol{P}$, they weight the contribution of the $\boldsymbol{P}$ injection into each $\boldsymbol{M}$ band on the correlation between $\boldsymbol{M}$ and the synthetic panchromatic image $\boldsymbol{J}$ to produce a set of high resolution intensity bands. The difference between these bands and $\boldsymbol{J}$ form the high-frequency information that are modulated with the gains vector ($\boldsymbol{g}_n$) and added onto the original spectral bands $\boldsymbol{M}$ in the usual CS framework (Eq. (5.5)).

Recently, Leung et al. (2014) proposed the Improved-Adaptive-IHS (IAIHS) that expands on the study of Rahmani et al. (2010) who proposed that the gains $\boldsymbol{g}_n$ be dependent on the edges of $\boldsymbol{P}$. Under the rationale that injecting the same amount of detail into different bands promotes spectral distortion, Leung et al. (2014) instead proposed using gains based on the spectral band specific edges, such that the ratios between the bands are preserved. Their algorithm excelled with scenes containing areas of vegetation, beating AIHS and standard multi-resolution based methods with respect to preserving correlation with $\boldsymbol{M}$.

5.1.2 Principal Component Analysis (PCA)

PCA is a multi-disciplinary tool used in many areas of science. It is a statistical technique that transforms a multivariate dataset of correlated variables into a new uncorrelated dataset that has been constructed from a linear combinations of the original variables. PCA rotates the original basis such that one of the axis (the first principal component) defines the vector direction of maximum statistical variance of the multivariate dataset. The remaining orthogonal axis point in deceasing directions of variance and it is the projection of the original data onto these new axis that provides the newly constructed uncorrelated dataset. For images, it creates an uncorrelated feature space that serves the hypothesis in pan-sharpening that the spatial information (from all the spectral bands) is contained within the first principal component, to the extent proportional to the correlation amongst $\boldsymbol{M}$ (Vivone et al., 2015) while the spectral information is accounted for by the remaining principal components (Pat S. Chavez and Kwarteng, 1989).

Mathematically, PCA is a matrix decomposition that operates on a multivariate data positive-definite covariance (unstandardised) or correlation (standardised - normalised variance) matrix. The use of the correlation matrix scales the axis so that the features receive a unit variance, this prevents certain features/statistical anomalous/noisy data from dominating the image because of their large values (Pohl and van Genderen, 1998) that one might otherwise see with a unnormalised covariance matrix. In addition it also removes the differences in dynamic range that might exist between the channels (Yésou et al., 1993).

Pan-sharpening with PCA requires projecting all the multispectral bands $\boldsymbol{M}$ onto the eigenvector associated with the largest eigenvalue (the first principal component), this new high spatial resolution image can be referred to as PC1 in keeping with the notation in the literature. Projecting onto the additional components can be referred to as PC2, PC3 etc. In the same vein as the IHS, now that the image as been decomposed into spectral and spatial components, substitution of PC1 with $\boldsymbol{P}$ can occur, after which the inverse transform produces the pan-sharpened image (Chavez et al., 1991).

The injection gains $\boldsymbol{g}_n$ per band has been stated by Vivone et al. (2015) to be defined by the first column in the eigenvector matrix $\boldsymbol{U}$.

It is well established that pan-sharpening from PCA is superior to IHS (Chavez et al., 1991) not only in that it is not limited to three spectral bands, but also in post fusion statistical evaluation. It does suffer from similar problems to that of the IHS method; the spectral characteristics of PC1 will never exactly match the spectral response of $\boldsymbol{P}$.

We know PCA based pan-sharpening provides high levels of spatial enhancement of $\boldsymbol{M}$ at the cost of altering their spectral characteristics (González-Audícana et al., 2004). In an attempt to fix this unwanted spectral distortion caused by standard PCA pan-sharpening, Shahdoosti and Ghassemian (2011) and Shahdoosti and Ghassemian (2016) proposed combining "spectral-PCA (SPCA)" with regular PCA. They rationalise that because of this high correlation, the difference in the global statistics between the bands corresponds with the spectral dispersion between the bands. Histogram matching $\boldsymbol{P}$ to each $\boldsymbol{M}$ band will provide a PC1 that contains the spectrally distorted characteristics, which if substituted with $\boldsymbol{M}$ (statistically matched to PC1) will provide an accurate spectral band-set to be used with the spatially enhanced band-set obtained via the regular PCA-CS based method. Combining SPCA with regular PCA can be done by low-pass filtering the original PCA pan-sharpened bands and adding it onto the the low-pass filtered SPCA bands.

5.1.3 Gram-Schmidt (GS)

The GS transform is a well-known technique used in linear algebra and multivariate statistics that uses a series of projections to orthogonalise a set of vectors or - similarly to PCA - to produce an uncorrelated basis (Fraleigh and Beauregard, 1995). The GS method for pan-sharpening was invented by Laben and Brover in 1998, patented by Kodak (Laben and Brower, 2000) and is apart of the software package ENVI-IDL, where there exist two modes of operation: GS1 and GS2. The GS1 mode starts from a pre-defined vector, which is chosen to be a mean-centred synthetic low resolution approximation of $\boldsymbol{P}$ (usually taken as an average for all bands, $\boldsymbol{I}$). Following GS orthogonalisation, pan-sharpening is accomplished by replacing $\boldsymbol{I}$ with $\boldsymbol{P}$ before the inverse transformation is performed to obtain the new RGB vectors. The GS technique is often referred to as a generalisation of PCA; the GS technique reduces to the standard PCA method when PC1 is used as $\boldsymbol{I}$ (Aiazzi et al., 2009). The GS2 mode uses a low-filtered version of the original panchromatic image as the starting vector instead of $\boldsymbol{I}$ used in GS1. Amro et al. (2011) reports that GS1 produces pan-sharpened images with outstanding spatial quality at the cost of introducing spectral distortions, whereas GS2 suffers from lower spatial enhancement but superior spectral preservation. The low resolution intensity image described by Aiazzi et al. (2007) has been applied to the GS2 mode as adaptive-GS (or GSA) where it was reported to have reduced spectral distortions over the standard GS mode, especially on vegetated areas.

The injection gains $\boldsymbol{g}_n$ has been given by Aiazzi et al. (2007) as $\frac{\mathrm{cov}(\boldsymbol{M}_n,\boldsymbol{I})}{\mathrm{var}(\boldsymbol{I})}$.

5.1.4 Intensity Modulation

Although generally labelled among methods featuring an RSC, the Brovey Transform (BT) has been cited (Alparone et al., 2004) as being based on the chromaticity transform (Gillespie et al., 1987). The BT seeks to preserve the chromatic properties of an image by modulating the magnitude of the RGB vectors by multiplying the original RGB by a high resolution image. This process is known as Intensity Modulation (IM) and in the case of the BT, it serves the purpose of spatial enhancement of three band images, by modulating the original intensity mapping by that of a higher resolution one. Assuming that $\boldsymbol{P}$ has the same mean as $\boldsymbol{I}$, the BT preserves the global mean of the RGB from construction of the modulating image; a ratio of a high resolution image to that of the intensity image. The BT can be considered to be a CS method. This has been shown by Vivone et al. (2015), where they substitute the gain: $\boldsymbol{g}_n = \frac{\boldsymbol{M}_n}{\boldsymbol{I}}$ into Eq. (5.5), which results in the famous BT equation

$$\boldsymbol{H}_n = \boldsymbol{M}_n \circ \frac{\boldsymbol{P}}{\boldsymbol{I}}, \tag{5.7}$$

where $\circ$ represents the Hadamard product; a matrix multiplication where the matrix elements at their row/column location are multiplied by the other matrix elements at the same row/column location. The denominator of Eq. (5.7) can either be (as we have shown) $\boldsymbol{I}$ or a constructed image $\boldsymbol{J}$ using the RSC approach. As such it can either be considered to be a CS or an RSC technique.

It has been reported that the BT provides excellent improvement in contrast and preserves the chromatic information yet suffers with respects to spectral distortion in the image (Vijayaraj et al., 2004). Many authors cite the reasons for the spectral distortion as being down to the difference in spectral range of the intensity (or modulation) image being different to spectral range covered by the colour composition (Chavez et al., 1991, Guo et al., 1997, Guo and Moore, 1998, Liu, 2000, Ranchin and Wald, 2000, Alparone et al., 2004).

The BT was developed to enhance contrast in features such as shadows, water and high reflectance areas. As it is prone to spectral distortion, it should not be used if preserving the original scene radiometry is important. It is good for producing RGB images with a higher degree of contrast (Fonseca et al., 2011).

The general form of the BT that operates on images with an arbitrary number (d) of spectral bands is obtained by swapping out $\boldsymbol{I}$ on the numerator with the total sum of the $\boldsymbol{M}$ bands (Pohl and van Genderen, 1998, Wiemker et al., 1998)

$$\boldsymbol{H}_k = \boldsymbol{M}_n \circ \frac{\boldsymbol{P}}{\sum\limits_n^d \boldsymbol{M}_n}. \tag{5.8}$$

Amro et al. (2011) outline the use of Eq. (5.8) where the panchromatic image is statistically matched to each spectral (upsampled to the same resolution as $\boldsymbol{P}$) band to produce the new pan-sharpened bands. Variations of the IM method have been published: Synthetic Variable Ratio method, proposed by Munechika et al. (1993), which replaces the denominator of Eq. (5.8) with $\boldsymbol{I}$ that is constructed using the weighted sum (Eq. (5.6)) of the upsampled $\boldsymbol{M}$ bands. Using reflectance spectra for five land and three different atmosphere types, a linear least squares optimisation was carried out to calculate the band weightings for the synthetic panchromatic image $\boldsymbol{J}$.

5.1.5 Synthetic Variable Ratio

Generating new bands based on ratio methods that utilise the RSC date back to the 1980s. Price (1987) produced a method of pan-sharpening that relied on the assumption that because of the high correlation between the green, red and panchromatic bands, the downsampled $\boldsymbol{P}$ ($\boldsymbol{P}^L$; equally weighted mean of a pixel neighbourhood) can express any of the $\boldsymbol{M}$ bands subject to a scaling and offset operation

$$\boldsymbol{M}_n^J = \alpha_n \boldsymbol{P}^L + \beta_n, \tag{5.9}$$

where $\boldsymbol{M}_n^J$ is the approximation of the original $\boldsymbol{M}_n$ spectral band. The weights α and β can be determined using a linear least squares optimisation, which when applied to $\boldsymbol{P}$, produces high resolution spectral bands $\boldsymbol{M}_n^H$ band estimations that possess local means (in a super pixel i.e. a 2×2 pixel block) that are the same as the registered pixel in the low resolution spectral band. The following pan-sharpened image is then calculated using

$$\boldsymbol{H}_n = \boldsymbol{M}_n \circ \frac{\boldsymbol{M}_n^H}{\boldsymbol{M}_n^{H,L}}, \tag{5.10}$$

where the numerator is essentially a histogram stretched $\boldsymbol{P}$ (for each band) and the denominator is its downsampled counterpart. The authors bring the images to the same size using a nearest neighbour interpolation. Following this, Munechika et al. (1993) proposed a method which replaces the denominator of Eq. (5.8) with a $\boldsymbol{J}$ that is constructed with Eq. (5.6) using upsampled $\boldsymbol{M}$ bands. To save on computation time, a linear least squares optimisation was used to compute the band weights for specific landscapes and viewing conditions. These were based on reflectance spectra for five and three different land and atmosphere types respectively. $\boldsymbol{P}$ was histogram matched to this $\boldsymbol{J}$ such that atmospheric and illumination differences between the two could be minimised.

In order to free up the user from having to decide on these fusion parameters, Zhang (1999) proposed calculating a high resolution $\boldsymbol{J}$ using the least squares approach on upsampled $\boldsymbol{M}$ bands. They forego the histogram matching process of $\boldsymbol{P}$ to $\boldsymbol{J}$ because on the high correlation of the Landsat Thematic Mapper bands with the SPOT panchromatic.

Wang et al. (2008) suggested calculating the weights based on the sensor response of the satellite, where they proposed a fast integral approximation that takes into account not just the areas under the band sensor spectrographs, but also the regions of low/zero sensitivity between them. Similarly, Chen et al. (2008) tackled the problem using the ratio of the integral of the spectral sensor characteristics to that of the overlap with the panchromatic spectral response. Instead of modulating their ratio with the spectral bands they used the luminosity component from CIE $L^*a^*b^*$ space, which required the addition of an RGB to CIE $L^*a^*b^*$ space transform and its inverse before producing the desired pan-sharpened image. This satellite dependent means of calculating the coefficients is popular as it speeds up the fusion process (the coefficients only need to be calculated once). The disadvantages to this approach are 1) prior knowledge of the satellite is required and 2) a different set of weights would need to be computed for platforms that have variations in their sensors.

5.2 High-Pass Frequency Injection

To prevent distortions of the low-frequency component that the RSC/CS can cause, High-Pass Filtering (HPF) of $\boldsymbol{P}$ was constructed on the basis of providing a edge map that only effects the high-frequency components of $\boldsymbol{M}$. Introduced by Schowengerdt (1980) and modified by

Chavez et al. (1991), the panchromatic image is convolved with a high-pass filter kernel; a process that removes the low-frequency information. Alternatively, a low-pass filtered version of the panchromatic image can be subtracted from the original to provide the high-frequency edge information.

One of the problems with this type of fusion is that the size of the filter kernel is not arbitrary however, Chavez et al. (1991) have suggested that the best size is approximately twice the size of the ratio of the spatial resolutions of the sensors. Wald et al. (1997) state that pan-sharpening methods based on injecting high-frequency components into resampled versions of the spectral data have demonstrated a superior performance and compared with many other pan-sharpening methods such as CS.

Liu (2000) apply a box filter kernel to $\boldsymbol{P}$ that is directly based on the size of its ratio with $\boldsymbol{M}$. Using a ratio, this low-pass filtered $\boldsymbol{P}$ is used to extract the high-frequency information from band dependent statistically matched $\boldsymbol{P}$ which are used as an IM image for $\boldsymbol{M}$.

Li et al. (2015) have applied a Gaussian filter ($\sigma = 3$) to decompose the $\boldsymbol{P}$ and $\boldsymbol{M}$ images into their high/low -frequency components. They create low resolution panchromatic images by adding the low-frequency parts of $\boldsymbol{P}$ onto the high-frequency components of the $\boldsymbol{M}$ bands. Using the SVR method, they take the ratio of the original $\boldsymbol{M}$ bands with these new pan-like spectral bands and modulating it with $\boldsymbol{P}$ provides their desired high resolution $\boldsymbol{H}$ bands. They reason that by taking this ratio, they can successfully cancel out the low-frequency (spectral) information contained in $\boldsymbol{P}$ and the spatial information in the original $\boldsymbol{M}$ bands which results in a sharpened image $\boldsymbol{H}$ with the spatial and spectral characteristics of the panchromatic and original multispectral image respectively.

5.3 Multi-Resolution Analysis (MRA)

Ranchin et al. (1996) proposed a model of panchromatic image fusion that operates on the principle that $\boldsymbol{M}$ is merely a downsampled (blurred) version of the desired high resolution image $\boldsymbol{H}$. Using the information in $\boldsymbol{P}$, the goal is to solve for this high resolution image $\boldsymbol{H}$, an inverse problem which the authors originally posed using a Multi-Resolution Analysis (Ranchin and Wald, 2000) framework. This concept has been referred to as ARSIS (Amélioation de la

Résolution Spatiale par Injection de Structures).

Overall the premise behind the MRA pan-sharpening is the same as the previous methods; to extract the spatial detail from $\boldsymbol{P}$ and inject it into the resampled $\boldsymbol{M}$ bands. However, the MRA class of pan-sharpening offers a more selective (compared to global methods at a single resolution) means of fusion. By decomposing the image to an ordered set of resolutions image manipulation can take place at the feature level (high-frequency components disappear at low resolutions leaving only the coarser structures).

5.3.1 Wavelets

Wavelet analysis originated in signal processing (Mallat, 1989) and has found applications in image pan-sharpening. The discrete wavelet transform (DWT) is a frequency domain transform that is localised in time and space making it ideal for extracting image details. It has become a popular tool for pan-sharpening for its ability to detect/inject high-frequency components into the $\boldsymbol{M}$ bands.

Mallat (1989) originally implemented the DWT using a recursive dyadic subsampling and decimation process (MWT). For images (Zhou et al., 1998), this MRA can be represents by a pyramid, where the base is the original image, with successive levels representing coarser and coarser resolutions. At each level in the pyramid, the DWT outputs three wavelet coefficient images which contain the high-frequency components lost between the changes in resolution in the three directions: horizontal, vertical and diagonal. When the DWT process is inverted, the original image can be reconstructed exactly from an approximation image and the wavelet coefficients.

Another classical implementation of the DWT is known as the 'à trous' (Holschneider and Tchamitchian, 1989) wavelet transform (ATWT). In contrast to MWT, the ATWT is undecimated making it shift invariant. The approximation images remain the same size despite being downsampled in resolution at each level. At each of these levels, the ATWT outputs a single image that contains the horizontal, vertical and diagonal spatial information that is lost between the levels. The original image can be recovered exactly by using the inverse transform and adding the wavelet plane. Unlike MWT which is critically subsampled, the ATWT is oversampled (which will contain redundancies). The omission of the decimation step allows an image to be decomposed into nearly disjointed bandpass channels in the spatial frequency domain, without losing the spatial connectivity (translation invariance property) of its high-pass details e.g, edges and textures. This property is fundamental because, for critically subsampled schemes, spatial distortions, typically ringing or aliasing effects may be present in the fused image which can be seen as blur of contours and textures (Stathaki, 2008). As a consequence of this, the ATWT is seen to perform better at the pan-sharpening problem than decimated approaches (Aiazzi et al., 2002). González-Audícana et al. (2005) have studied in detail the differences between MWT and the ATWT and how to apply them to the pan-sharpening problem. Generally speaking, image fusion based on the wavelet transform (WT) preserves the spectral characteristics of fused satellite images over that of the IHS method (Yocky, 1996, Zhou et al., 1998, Choi et al., 2005), but provides much less spatial information (Gharbia et al., 2014). The problem with wavelet based fusion methods is that they are not efficient enough to quickly merge massive volumes of data from recent satellite images because of the higher computational complexity (Choi, 2006).

5.4 Hybrid Methods

CS/RSC pan-sharpening methods produce images with good spatial enhancement at the cost of spectral distortion. MRA pan-sharpening excel at preserving the spectral properties, and in some cases can even match the spatial enhancement of the CS i.e. the undecimated DWT. The problem of these frequency domain methods is in their computational complexity. Hybrid methods that attempt to incorporate the strengths of CS and MRA methods have become the direction to address the pan-sharpening problem. The general form of the hybrid method is to create a new spatial component i.e. a new intensity image in the case of IHS by injecting the edge information of $\boldsymbol{P}$ into the original $\boldsymbol{I}$.

Based on the IHS transform, Núñez et al. (1999) uses the ATWT to simply add the high resolution information from $\boldsymbol{P}$ directly onto the intensity image $\boldsymbol{I}$.

Hong and Zhang (2003) and Zhang and Hong (2005) investigate how different IHS transforms affect the low-frequency components of the $\boldsymbol{I}$ mappings once they have been substituted into $\boldsymbol{P}$ for wavelet fusion.

Ling et al. (2007) uses partial replacement of $\boldsymbol{I}$ based on fast Fourier transform (FFT) filtering. Low-pass filtering $\boldsymbol{I}$ and then substituting it with the complimentary high-pass information from $\boldsymbol{P}$ creates the new intensity image. Similarly, Mitianoudis et al. (2010) apply coefficient substitution of $\boldsymbol{P}$ with $\boldsymbol{I}$ in the wavelet domain using a decimated WT. To reduce the complexity, they forego the initial upsampling step on the rationale that downsampling occurs during the WT and that all that is required to maintain a resolution between $\boldsymbol{P}$ and $\boldsymbol{I}$, is to downsample $\boldsymbol{P}$ until it matches the resolution of the $\boldsymbol{I}$.

Shettigara (1992) applied the additive wavelet method to PCA, where they add the spatial detail from $\boldsymbol{P}$ that is is missing from $\boldsymbol{M}$ to the first principle component mapping PC1.

González-Audícana et al. (2004) investigated the quality of using decimated/undecimated WT based mergers with IHS and PCA techniques and found them to give superior fusion results to their stand-alone means of pan-sharpening, with the best results coming from the undecimated WT.

An extension to the wavelet concept, the Curvelet transform (Ma and Plonka, 2010) boasts superior representation of directional edges and has been applied to IHS pan-sharpening (Valizadeh and Ghassemian, 2012) where it produced superior spectral properties to that of IHS and ATWT.

The Radon transform has been applied to pan-sharpening as an IHS hybrid approach. The Radon transform is an integral transform that has found use in medical imaging. As with all IHS hybrid approaches, the $\boldsymbol{I}$ component is isolated and transformed. Sujitha et al. (2013) substitute the intensity images Radon coefficients with the $\boldsymbol{P}$ Radon coefficients using the maximum fusion rule, which means that the larger of the two is selected for the fused image. The inverse transform provides the final intensity mapping for the pan-sharpened image. Although this approach was shown to provide higher information and lower spectral errors than other techniques, it unfortunately requires two computationally complex transforms.

Adopting a similar fusion method to González-Audícana et al. (2004), Shah et al. (2008) use the Contourlet transform (CT) with a hybrid PCA approach. The CT is based on the Laplacian pyramid decomposition with direction filter banks applied on each bandpass sub-band. Instead of using PC1, they use the PC that has the highest correlation with $\boldsymbol{P}$ (which they call adaptive PCA). PC substitutions of the detail Contourlet coefficients are made with $\boldsymbol{P}$ before the inverse transform provides the final pan-sharpened image. Evaluation of the post-fused images using global metrics showed superior results in nearly all cases against the wavelet PCA approach (González-Audícana et al., 2004).

The non-subsampled Contourlet transform (NSCT) has been used to separate the high frequency information from $\boldsymbol{P}$ to inject into $\boldsymbol{M}$ for pan-sharpening (Mahyari and Yazdi, 2011). The NSCT has been previously applied to denoising, enhancement and contour detection (da Cunha et al., 2006). As with the ATWT, non-subsampling makes the NSCT shift invariant. Additionally, it allows for redundancies in the basis functions which results in a lower complexity than the CT. So as not to introduce noise into the pan-sharpened image, the authors forego using the maximum fusion rule and instead inject the non-subsampled Contourlet coefficients from $\boldsymbol{P}$ into $\boldsymbol{M}$ on the condition that their correlation be similar (as defined by a threshold).

Using calculus of variations, Zhou et al. (2013) construct a high resolution $\boldsymbol{I}$ that contains the same high-frequency components as that of $\boldsymbol{P}$. In order to preserve the spectral properties, partial differential equations that describe the gradient information in the fused image $\boldsymbol{J}$ are constructed such that its low-pass filtered version approximates the original intensity mapping $\boldsymbol{I}$. The partial differential equations are solved using the gradient descent method. Once $\boldsymbol{J}$ has been constructed it is substituted into the IHS method and converted back to RGB space.

Kaplan et al. (2012) extract the image detail of $\boldsymbol{P}$ using the bilateral filter before modulation and addition onto the $\boldsymbol{M}$ bands. It has been reported that the bilateral filter can sometimes produce gradient reversal artefacts, a problem that was fixed by the creation of the guided filter (He et al., 2013). As with the joint bilateral filter, the guided filter uses a guiding image with which to approximate to a filtered output. In pan-sharpening, $\boldsymbol{I}$ would be guided by the high-pass filtered $\boldsymbol{P}$ (Jameel et al., 2016) to produce a new intensity mapping for IHS technique (Jameel et al., 2016, Yang et al., 2016).

Using the injection model (Eq. (5.5)), Shahdoosti and Ghassemian (2015) computed $\boldsymbol{J}$ by using the appropriate filter 'h' on $\boldsymbol{P}$. This filter must produce the pan-sharpened spectral bands ($\boldsymbol{H}_n$) that minimise the sum of the mean square errors of $\boldsymbol{M}_n$ and their histogram matched $\boldsymbol{P}_n$. This constrained optimisation is solved using the Lagrangian multiplier method which has a low computation complexity that is limited by computing correlation matrices. The final pan-sharpened images were competitive against Bayesian, improved-IHS, sparsity, non-subsampled Contourlet and other filter based approaches.

5.5 Analysis of Pan-Sharpening by Linear Mappings

Recall the general scheme for CS

$$\boldsymbol{H}_n = \boldsymbol{M}_n + \boldsymbol{g}_n(\boldsymbol{P} - \boldsymbol{J}), \tag{5.11}$$

where $\boldsymbol{J}$ is our greyscale created from a weighted sum of the spectral bands

$$\boldsymbol{J} = \sum_{1}^{d} \omega_n \boldsymbol{M}_n \tag{5.12}$$

and d is the total number of spectral bands. Various calculations of the spectral band weights $\omega_n(\geq 0)$ and the injection gains $\boldsymbol{g}_n$ have been proposed. Pre-defined static weights are shown in Table. 5.1, and are used in the classic IHS transform, the GIHS (Tu et al., 2001) and the FGIHS (Tu et al., 2004). The first, while being fast, introduced large spectral errors and is limited to only three bands. The later two aim to reduce this by generalising the intensity mapping by allowing for the inclusion of the $\boldsymbol{NIR}$ band thus increasing the correlation with $\boldsymbol{P}$. To reduce spectral distortion, redundancies due to the spectral overlap in the blue and green bands were taken into account by altering the weights, which were calculated based on empirical data taken from the IKONOS satellite (Tu et al., 2004). Unfortunately these pre-defined weights limit the application to imagery of landmasses dominated by vegetation.

To increase the utility of previous IHS methods, image dependent/adaptive methods of determining the band weights were developed. Adopting the method found in colour image sharpening, $\boldsymbol{M}$ band weights are calculated to produce $\boldsymbol{J}$ that best approximates the $\boldsymbol{P}$ in a least squares sense. This is based on the assumption that $\boldsymbol{J}$ is a blurred version of $\boldsymbol{P}$ which when applied to the second term of Eq. (5.5) results in an edge map that contains the high-frequency components missing from the spectral bands. These components are injected into $\mathbf{M}$ by arithmetic addition. Computing $\boldsymbol{J}$ has been carried out by linear regression (Xu et al., 2008, Choi et al., 2011) or multivariate calculus (Rahmani et al., 2010, Zhou et al., 2013).

The assumption that $\boldsymbol{J}$ is a low-frequency filtered version of the $\boldsymbol{P}$ wrongly assumes that there is a near perfect construction of the panchromatic low-frequency components from the

Table 5.1: Summary of static channel weights and injection gains.

Method	$\omega_{\boldsymbol{B}}$	$\omega_{\boldsymbol{G}}$	$\omega_{\boldsymbol{R}}$	$\omega_{\boldsymbol{NIR}}$	g_n
IHS	1/3	1/3	1/3	0	1
GIHS	1/4	1/4	1/4	1/4	1
FGIHS	1/12	1/4	1/3	1/3	1

linear combination of spectral bands. To address this flaw, it is common practice to use the injection gains $\boldsymbol{g}_n$ as an edge map operator ($\varepsilon_{\mathbf{A}}$) that extracts the high-frequency components. Rahmani et al. (2010) use an exponential edge detector on their panchromatic image (AIHS)

$$\varepsilon_{\boldsymbol{A}} = \exp\left(-\frac{\lambda}{|\nabla \boldsymbol{A}|^4 + \epsilon}\right), \tag{5.13}$$

where $\boldsymbol{A}$ is the image whose edges are being extracted ($\varepsilon_{\boldsymbol{P}}$ for $\boldsymbol{P}$ and $\varepsilon_{\boldsymbol{M}}$ for $\boldsymbol{M}$), λ being a user defined edge tolerance parameter and ϵ is a small constant to avoid division by zero. The AIHS method has the simplest means of extracting and injecting edges. It merely extracts from the panchromatic image and adds them into the spectral bands, which is the same for each band and will undesirably change the spectral bands vector direction in the final fused image, thus introducing spectral distortion (Tu et al., 2001). To reduce the spectral distortion in areas of the image where edge injection has occured, Leung et al. (2014) believes that the injection into each band should be different and suggest a user defined (β) parameter that provides for a compromise between $\boldsymbol{P}$ injection and $\boldsymbol{M}$ edge enhancement for each individual spectral band (IAIHS) that

Table 5.2: The methods of calculating weights and injection gains for the CS schemes.

Method	ω	g_n
BT	$\min\limits_{\omega} \lvert\lvert \boldsymbol{P} - \boldsymbol{J} \rvert\rvert^2$	$\frac{\boldsymbol{M}_n}{\boldsymbol{I}}$
AIHS	$\min\limits_{\omega} \lvert\lvert \boldsymbol{P} - \boldsymbol{J} \rvert\rvert^2$	$\varepsilon_{\boldsymbol{P}}$
IAIHS	$\min\limits_{\omega} \lvert\lvert \boldsymbol{P} - \boldsymbol{J} \rvert\rvert^2$	$\frac{\boldsymbol{M}_n}{\frac{1}{d}\sum_1^d \boldsymbol{M}_n} \left(\rho\varepsilon_{\boldsymbol{P}} + (1-\rho)\varepsilon_{\boldsymbol{M}}\right)$
PRACS	$\min\limits_{\omega} \lvert\lvert \boldsymbol{P} - \boldsymbol{J} \rvert\rvert^2$	$\beta \mathrm{CC}(\boldsymbol{J}, \boldsymbol{M}_n) \frac{\mathrm{std}(\boldsymbol{M}_n)}{\frac{1}{d}\sum_1^d \mathrm{std}(\boldsymbol{M}_n)} \boldsymbol{L}_n$
WLS-IHS	$\min\limits_{\omega} \lvert\lvert \boldsymbol{P}^H - \boldsymbol{J}^H \rvert\rvert^2$	$\left[\min\limits_{\boldsymbol{D}_n} \quad \left\lVert \boldsymbol{P}^H - \sum_1^d \omega_n \left(\boldsymbol{M}_n^H + \boldsymbol{D}_n\right)\right\rVert^2\right] \frac{1}{\boldsymbol{D}_n}$
MGF-IAIHS	$\min\limits_{\omega} \lvert\lvert \boldsymbol{P} - \boldsymbol{J} \rvert\rvert^2$	$\frac{\boldsymbol{M}_n}{\frac{1}{d}\sum_1^d \boldsymbol{M}_n} \left(\rho_n\varepsilon_{\boldsymbol{P}} + (1-\rho_n)\varepsilon_{\boldsymbol{M}}\right)$

is modulated by the normalised spectral band pixel ratio. This alleviates the problem of over compensating already salient edges that appear in both the spectral and panchromatic bands, a problem that can occur with the AIHS method. The disadvantage is that this applies to all spectral bands, which while not necessarily a problem for many edges (due to the large degree of correlation between the spectral bands), can still be a problem for some where edges appear in $\boldsymbol{P}$ and not in certain $\boldsymbol{M}$ bands (Thomas et al., 2008).

Yang et al. (2016) (MGF-IAIHS) seeks to address this problem by proposing a band dependent calculation of the ρ parameter that satisfies the optimisation that the edge filtered $\boldsymbol{P}$ is a linear combination of the edge filtered $\boldsymbol{M}$ bands

$$\min_{\rho} \quad \left\| \varepsilon_{\boldsymbol{P}} - \sum_{n=1}^{d} \rho_i \varepsilon_{\boldsymbol{M}} \right\|^2 , \tag{5.14}$$

which is then used to calculate the band dependent gains in a similar manner to IAIHS. For approximation of the panchromatic image, $\boldsymbol{J}$ is used as the guiding image for the multi-layered guided filter (MGF-IAIHS) $\boldsymbol{P}_{\mathbf{D}}^{j} = GF^{j-1}(\boldsymbol{P}) - GF^{j}(\boldsymbol{P}, \boldsymbol{J})$, where the sum of the '$j$' layers provides the edge map in Eq. (5.5)

$$\boldsymbol{H}_n = \boldsymbol{M}_n + \boldsymbol{g}_n \sum_{j=1}^{K} \boldsymbol{P}_{\boldsymbol{D}}^{j}. \tag{5.15}$$

Song et al. (2016) propose a weighted least squares filter (WLS-IHS) to compute the high-frequency/edge maps for both the panchromatic ($\boldsymbol{P}^H$) and spectral bands ($\boldsymbol{M}_n^H$). Using a linear combination of $\boldsymbol{M}_n^H$ that best approximates $\boldsymbol{P}^H$ provides $\boldsymbol{J}^H$, which is then used to calculate the edge map $\boldsymbol{D}_n = \boldsymbol{P}^H - \boldsymbol{J}^H$. Specifically, ω_n is calculated by

$$\omega_n = \left(\left(\boldsymbol{M}_n^H\right)^T \boldsymbol{M}_n^H \right)^{-1} \left(\boldsymbol{M}_n^H\right)^T \boldsymbol{P}^H. \tag{5.16}$$

Instead of using an edge map operator, Choi et al. (2011) extract the high-frequency components based upon a partial replacement approach. The authors propose constructing a panchromatic image for each spectral band that is based on the correlation with $\boldsymbol{J}$

$$\boldsymbol{P}_n = \mathrm{CC}(\boldsymbol{J}, \boldsymbol{M}_n)\boldsymbol{P} + (1 - \mathrm{CC}(\boldsymbol{J}, \boldsymbol{M}_n))\boldsymbol{M}_n, \tag{5.17}$$

where $\mathrm{CC}(\boldsymbol{J}, \boldsymbol{M}_n)$ is the correlation coefficient between the approximated panchromatic and the n^{th} spectral band. $\boldsymbol{P}_n$ is then used as a set of target images for another linear regression of $\boldsymbol{M}$ to produce band specific synthetic $\boldsymbol{J}_n$. The injection gains are calculated by in two parts, a global and local computation. The global component starts by calculating a band specific constant that is produced by normalising the standard deviation of $\boldsymbol{M}$ with the mean standard deviation across all bands and is finished by the product of the bands correlation with $\boldsymbol{J}_n$. Further to this, a tuning parameter β is utilised to stabilise the dynamic range of the final image. The pixel and band specific adaptive factor $\boldsymbol{L}_n$ serves the purpose of reducing what the authors define as the spectral instability error between the synthetic component image and the spectral band and is calculated by

$$\boldsymbol{L}_n = \left(1 - \left|1 - \mathrm{CC}(\boldsymbol{J}, \boldsymbol{M}_n)\frac{\boldsymbol{M}_n}{\boldsymbol{J}_n}\right|\right). \tag{5.18}$$

5.6 Conclusion

In this chapter we introduced the component substitution (CS) method of pan-sharpening. Although this collection of methods has seen great success enhancing the spatial resolution of satellite imagery, it is known for distorting the spectral characteristics of the original multiband image.

The most popular CS method is based on the IHS transform where a greyscale $\boldsymbol{J}$ is created by linear mapping the spectral bands such that it has a maximum correlation with panchromatic band $\boldsymbol{P}$. This method is based on the assumption that $\boldsymbol{J}$ is a low-pass approximation of $\boldsymbol{P}$ and that one can extract its high-frequency information by taking their difference. The validity of this assumption is limited by the correlation of $\boldsymbol{J}$ with $\boldsymbol{P}$ which in turn is affected by the spectral overlap between the $\boldsymbol{M}$ and the $\boldsymbol{P}$ sensors.

The problem that researchers have since tried to address is how they go about injecting this information into the spectral bands without distorting the low-frequency components of the original image. At additional computational cost, hybrid methods based on the IHS transform have been created that rely on a) edge-detecting masks and b) multi-resolution analysis, each with their own unique way of extracting high-frequency information from $\boldsymbol{P}$ and injecting it into $\boldsymbol{I}$.

Despite the advent of these newer hybrid methods, the low complexity of CS based methods keeps it of interest to the research community. In the next chapter we present a novel adaptive method of constructing $\boldsymbol{J}$ that has maximum correlation with $\boldsymbol{P}$ subject to preserving the global statistics of $\boldsymbol{I}$.

Chapter 6

Multispectral Panchromatic Image Sharpening: Linear Mappings

The review in the last chapter introduced the popular component substitution injection scheme. In its simplest form, this method is faced with two challenges 1) how the synthetic panchromatic image $\boldsymbol{J}$ is computed and 2) how the edge map is injected into the spectral bands. The standard method to compute $\boldsymbol{J}$ is by linear regression. Ideally this would produce a perfect mapping of $\boldsymbol{J} = \boldsymbol{P}$. In this instance the correlation would be equal to 1 and the global statistics of both mappings would be identical. Unfortunately there is no known linear relationship between the $\boldsymbol{M}$ bands and $\boldsymbol{P}$ and the best we can hope for is a close approximation.

In this chapter we shall adopt a different approach to the problem. Bearing in mind the importance of correlation between our greyscale $\boldsymbol{J}$ with $\boldsymbol{P}$, we instead prioritise preserving the mean and variance of our mapping. The rationale behind this is that the ideal $\boldsymbol{J}$ would be equal to a $\boldsymbol{P}$ that had been statistically matched to $\boldsymbol{I}$. Furthermore, we impose that $\boldsymbol{J}$ be a real normalised image and that by introducing a constraint on variance we shall further reduce our solution space where the ideal mapping band weights would reside (if it existed). Lastly, from this reduced space, we introduce a method of computing $\boldsymbol{J}$ that, with respect to $\boldsymbol{I}$, will give a greater correlation with $\boldsymbol{P}$.

This chapter builds directly from our work in Chapter 3 and shall be structured accordingly. First we shall show how we can produce the closest approximation to the panchromatic image using a least squares optimisation. Secondly we shall show that by natural extension of our vector space we can produce a brightness preserving multiband-to-greyscale conversion and how using multivariate calculus we can produce the optimum $\boldsymbol{J}$ subject to preserving image mean. Lastly we shall introduce a novel constraint on variance which confines our solution space to lie on a elliptical line segment and how we can enhance our correlation subject to this constraint.

6.1 Linear Mapping for Maximum Correlation

If there was a linear relationship between the spectral bands $\boldsymbol{M}$ and the panchromatic image $\boldsymbol{P}$ then

$$\boldsymbol{P} = \boldsymbol{M}\boldsymbol{w}, \tag{6.1}$$

where $\boldsymbol{M}$ (dimensions $[l, w, d]$) and $\boldsymbol{P}$ ($[l, w, 1]$) have been arranged into an array of $[l \times w, d]$ and a single (l×w,1) vector.

As there is no known linear relationship between the spectral bands and the panchromatic image, the best we can achieve is to maximise the correlation between $\boldsymbol{J}$ and $\boldsymbol{P}$ by minimising the sum of the squares

$$\min_{\boldsymbol{w}} \|\boldsymbol{M}\boldsymbol{w} - \boldsymbol{P}\|^2 = \boldsymbol{w}^T\boldsymbol{M}^T\boldsymbol{M}\boldsymbol{w} - 2\boldsymbol{w}^T\boldsymbol{M}^T\boldsymbol{P} + \boldsymbol{P}^T\boldsymbol{P}. \tag{6.2}$$

We know from Finlayson and Matheson (2012) that in using multivariate calculus we can show that positive-definite quadratic forms possess a single minima. Differentiating with respect to $\boldsymbol{w}^T$ and rearranging for the minima gives

$$\boldsymbol{w}_{cor} = (2\boldsymbol{R})^{-1} 2\boldsymbol{M}^T\boldsymbol{P}, \tag{6.3}$$

where $\boldsymbol{R} = \boldsymbol{M}^T\boldsymbol{M}$ and $\boldsymbol{w}_{cor}$ is the vector in $\mathbb{R}^4$ that contains the band weights to produce $\boldsymbol{J}$ that has the maximum correlation with $\boldsymbol{P}$. As there are no constraints on this optimisation, it would be an almost certainty that $\boldsymbol{J}$ would have pixel values that are outside the viewing range and have different global statistics to $\boldsymbol{I}$.

To address the former we could, in a similar manner to that used in Chapter 3, produce a convex set Q in $\mathbb{R}^4$ from an intersection of half-spaces. If our $\boldsymbol{w}_{cor}$ were to exist outside of this set then we would need to compute the point on its hull that would produce $\boldsymbol{J}$ with the greatest correlation with $\boldsymbol{P}$. This is undesirable as it would involve computing a convex hull in 4 dimensions thus increasing the complexity of our solution.

Instead we choose to truncate the vector space by adding a constraint on the statistical mean of our mapping. In the next section we shall describe how we can reduce the dimensionality of our vector space by 2 by introducing constraints on image energy and mean.

6.2 Multispectral Brightness Preserving Greyscale Conversion

In Chapter 3 we showed that by linear mapping it was possible to create a greyscale that possessed the same average brightness as the original colour image. We shall now show that the addition of an extra dimension to our vector space (V) from $\mathbb{R}^3$ to $\mathbb{R}^4$ we can accommodate the $\boldsymbol{NIR}$ band to the brightness preserving multiband-to-greyscale.

Previously we defined the brightness of a pixel for three band colour image to be the arithmetic sum of the RGB triplet. Extending this to account for the $\boldsymbol{NIR}$

$$\boldsymbol{I} = \frac{\boldsymbol{R} + \boldsymbol{G} + \boldsymbol{B} + \boldsymbol{NIR}}{4}, \tag{6.4}$$

which is the intensity mapping used in the GIHS method (Tu et al., 2001). Preserving this averaging of the bands is a simple alteration of our previous equality constraint from Chapter 3 (Eq.(3.7))

$$B = \left\{ \boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T \boldsymbol{\mu} = \boldsymbol{u}^T \frac{\boldsymbol{\mu}}{4} \right\} \subseteq \mathbb{R}^4, \tag{6.5}$$

which we also apply to the energy

$$E = \{ \boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T \boldsymbol{u} = 1 \} \subseteq \mathbb{R}^4 \tag{6.6}$$

and

$$\boldsymbol{w}^T = [\alpha \quad \beta \quad \gamma \quad \epsilon], \quad \boldsymbol{u}^T = [1 \quad 1 \quad 1 \quad 1], \quad \boldsymbol{\mu}^T = [\boldsymbol{\mu}_R \quad \boldsymbol{\mu}_G \quad \boldsymbol{\mu}_B \quad \boldsymbol{\mu}_{NIR}], \tag{6.7}$$

with $L = B \cap E$; a space that has two 4-dimensional basis vectors $\boldsymbol{N}$ in $\mathbb{R}^4$

$$\boldsymbol{N} = \text{Null}\left(\begin{bmatrix} \boldsymbol{u}^T \\ \boldsymbol{\mu}^T \end{bmatrix}\right). \tag{6.8}$$

To ensure no pixel clipping in our greyscale we again pose two inequalities on each pixel ($\boldsymbol{h}_i$)

$$P_i = \{\boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T\boldsymbol{h}_i = k \quad | \quad i = 1, 2, ..., 2n, \quad 0 \leq k \leq 1\}. \tag{6.9}$$

Solving for the intersection of halfspaces P_i in L (described in Chapter 3) gives us a 2-dimensional convex set Q and this defines all possible band weightings that will give us a greyscale $\boldsymbol{J}$ that preserves the mean and energy of our multiband image. Using multivariate calculus, we shall now show how we can obtain the band weights for $\boldsymbol{J}$ that has a maximum correlation with $\boldsymbol{P}$ subject to the constraints defined by Eq. (6.5) and (6.6).

6.2.1 Constrained Maximum Correlation

In Section 6.1, we solved for the solution of maximum correlation using a least squares optimisation. Using the Lagrangian Multiplier method (Jordan and Smith, 1997) we can produce the solution for maximum variance subject to our equality constraints. Our optimisation becomes

$$\min_{\boldsymbol{w}} \|\boldsymbol{M}\boldsymbol{w} - \boldsymbol{P}\|^2 \quad \text{s.t.} \quad \boldsymbol{w} \in L. \tag{6.10}$$

Expanding out

$$L(\boldsymbol{w}, \boldsymbol{\lambda}) = \boldsymbol{w}^T\boldsymbol{M}^T\boldsymbol{M}\boldsymbol{w} - 2\boldsymbol{w}^T\boldsymbol{M}^T\boldsymbol{P} + \boldsymbol{P}^T\boldsymbol{P} - \lambda_1\left(\boldsymbol{w}^T\boldsymbol{u} - 1\right) - \lambda_2\left(\boldsymbol{w}^T\boldsymbol{\mu} - \boldsymbol{u}^T\frac{\boldsymbol{\mu}}{4}\right), \tag{6.11}$$

where $\boldsymbol{\lambda}$ is a vector that contains the Lagrangian Multipliers λ_1 and λ_2. Continuing with the

Lagrangian Multiplier method, we differentiate Eq. (6.11) with respect to each variable

$$\begin{aligned}\frac{\partial L}{\partial \boldsymbol{w}^T} &= 2\boldsymbol{R}\boldsymbol{w} - 2\boldsymbol{M}^T\boldsymbol{P} - \lambda_1\boldsymbol{u} - \lambda_2\boldsymbol{\mu} = 0,\\ \frac{\partial L}{\partial \lambda_1} &= \boldsymbol{w}^T\boldsymbol{u} - 1 = 0,\\ \frac{\partial L}{\partial \lambda_2} &= \boldsymbol{w}^T\boldsymbol{\mu} - \boldsymbol{u}^T\frac{\boldsymbol{\mu}}{4} = 0,\end{aligned} \tag{6.12}$$

which gives us a system of linear equations which we can solve using the methods of linear algebra

$$\begin{bmatrix}\boldsymbol{w}_{cor}\\ \boldsymbol{\lambda}\end{bmatrix} = \begin{bmatrix}2\boldsymbol{R} & -\boldsymbol{A}^T\\ \boldsymbol{A} & 0\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{M}\\ \boldsymbol{k}\end{bmatrix}, \tag{6.13}$$

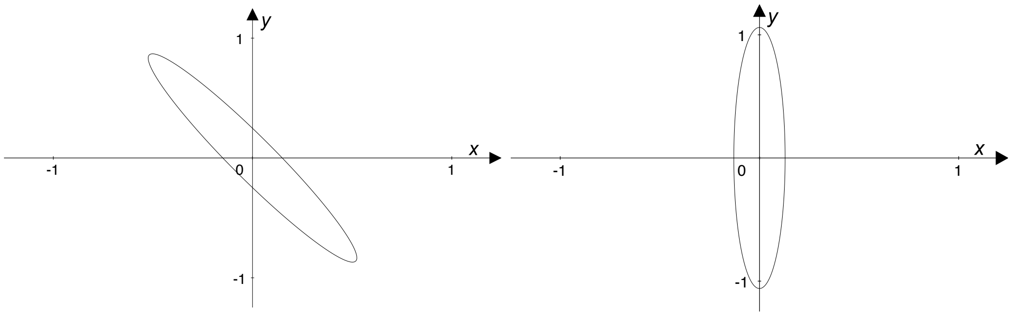
where $\boldsymbol{w}_{cor}$ is the vector that defines the solution for maximum correlation in $\mathbb{R}^4$ subject to L, $\boldsymbol{\lambda}$ the vector containing the Lagrangian Multipliers, $\boldsymbol{M}$ is a vector from $2\boldsymbol{M}^T\boldsymbol{P}$ and lastly $\boldsymbol{k}$ and $\boldsymbol{A}$ a vector and matrix that contain the constants and coefficients of our equality constraints respectively.

As in Section 6.1, our $\boldsymbol{w}_{cor}$ could produce band weights for an image that was beyond the dynamic range of our display and that we would need to find the closest correlation vector on the 2-dimensional convex set Q. In the next section we shall instead propose that by adding an additional constraint on the variance of our $\boldsymbol{J}$, we can reduce our solution space down to a single dimension, thereby simplifying the process of finding the point of nearest correlation.

6.3 Positive-Definite Quadratic Form and the Hyper-Ellipse

To satisfy a complete statistical matching operation during our linear mapping process we must produce an additional constraint that preserves the statistical variance. Recall that our variance is defined by our quadratic form $K = \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}$, and this describes the equation of a hyper-ellipse in non-homegeneous coordinates. For ease of description we shall refer to a 2-dimensional ellipse.

Performing singular value decomposition (SVD) on our covariance matrix ($\boldsymbol{\Sigma}$) decomposes it into a matrix of eigenvectors ($\boldsymbol{U}$) and eigenvalues ($\boldsymbol{S}$). The $\boldsymbol{S}$ matrix contains zeros with the



(a) $\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} = 1$ (b) $\boldsymbol{w}^T\boldsymbol{S}\boldsymbol{w} = 1$

Figure 6.1: (a) Ellipse that is described by the positive-definite quadratic form. (b) Rotated ellipse using the eigenvector basis.

eigenvalues ($\lambda_1 ... \lambda_d$) in descending order on the diagonal

$$\boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{U}^T. \tag{6.14}$$

The presence of off-diagonal zeros in $\boldsymbol{S}$ and the orthogonal property of $\boldsymbol{U}$ means that if we rearrange Eq. (6.14) for $\boldsymbol{S}$

$$\boldsymbol{S} = \boldsymbol{U}^T\boldsymbol{\Sigma}\boldsymbol{U} \tag{6.15}$$

and use it as our new covariance matrix in our quadratic form

$$\boldsymbol{w}^T\boldsymbol{S}\boldsymbol{w} = \boldsymbol{w}^T\boldsymbol{U}^T\boldsymbol{\Sigma}\boldsymbol{U}\boldsymbol{w}, \tag{6.16}$$

then we have rotated $\boldsymbol{\Sigma}$ and our quadratic form now describes the standard equation of an ellipse (see Figure. 6.1).

The major $a = \sqrt{K/\lambda_2}$ and minor $b = \sqrt{K/\lambda_1}$ semi-axis of our ellipse can now be calculated from $\boldsymbol{S}$ and K as $\sqrt{K}\sqrt{\boldsymbol{S}^{-1}}$. As we saw from Chapter 3, as we increase the Euclidean magnitude of $\boldsymbol{w}$ the variance K increases. Geometrically, this equates to our ellipse increasing in area while maintaining its aspect ratio. In our original frame of reference, the semi-major/minor axis of this ellipse can now be calculated from

$$\boldsymbol{M} = \sqrt{K}\boldsymbol{U}\sqrt{\boldsymbol{S}}^{-1}, \tag{6.17}$$

where the rows of $\boldsymbol{M}$ define the axis-vectors in our vector space. For our purposes we wish to preserve the variance of our mapping subject to equality constraints. Fundamentally this is a problem of solving for hyper-plane/hyper-ellipse intersections. We shall now describe how we go about solving this problem.

6.4 Multispectral Brightness and Contrast Preserving Greyscale Conversion

For a given variance K_p, we define our hyper-ellipse in $\mathbb{R}^4$ as the set S of all the vectors $\boldsymbol{w}$ that satisfy $\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} = K_p$ (see Figure. 6.2(a) for an analogue in one fewer dimension)

$$S = \{\boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} = K_p\} \subseteq \mathbb{R}^4. \tag{6.18}$$

Using the subscript o, let us also define E_o and B_o, the origins that have been translated from E and B

$$E_o = \{\boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T\boldsymbol{u} = 0\} \subseteq \mathbb{R}^4, \tag{6.19}$$

$$B_o = \{\boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T\boldsymbol{\mu} = 0\} \subseteq \mathbb{R}^4. \tag{6.20}$$

Again, $L = E \cap B$ defines the set from the intersection of two 3-dimensional hyper-planes that produces a 2-dimensional plane in $\mathbb{R}^4$ (see Figure. 6.2(b)). Similarly $L_o = E_o \cap B_o$ and this represents L as a subspace that passes through the origin of $\mathbb{R}^4$. Let us define the projection of all vectors $\boldsymbol{w} \in V$ onto L_o using the function

$$F : V \to L_o; \boldsymbol{w} \mapsto \boldsymbol{N}^T\boldsymbol{w}, \tag{6.21}$$

where the two columns of the matrix $\boldsymbol{N}$ represent the 4-dimensional basis vectors of L_o in $\mathbb{R}^4$ and can be computed by

$$\boldsymbol{N} = \text{Null}\left(\begin{bmatrix} \boldsymbol{u}^T \\ \boldsymbol{\mu}^T \end{bmatrix}\right). \tag{6.22}$$

The vectors $\boldsymbol{g}$ represent all the vectors in L_o and can be defined by $\boldsymbol{g} = F(\boldsymbol{w})$. Likewise we can define all vectors $\boldsymbol{g}$ using the basis of our original vector space V with $\boldsymbol{w} = F'(\boldsymbol{g})$

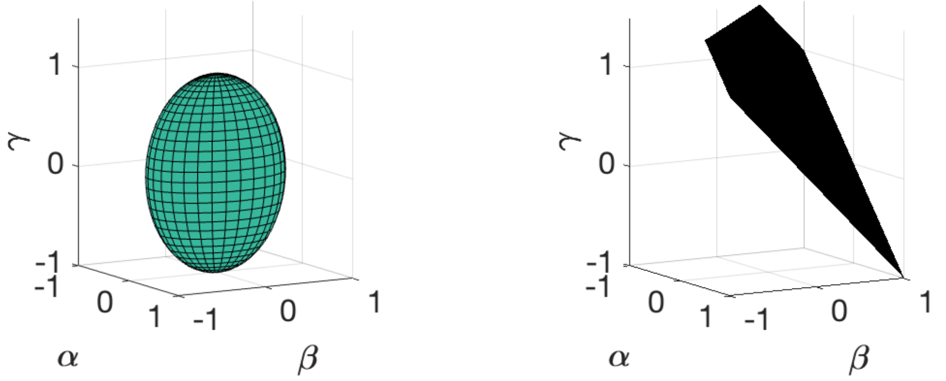
$$F' : L_o \to V; \boldsymbol{g} \mapsto \boldsymbol{N}\boldsymbol{g}. \tag{6.23}$$

$W = L \cap S$ defines the set that is the intersection of the 2-dimensional plane L with the 4-dimensional hyper-ellipse S, and this produces a 2-dimensional ellipse W (see Figure. 6.2(c)-(d)) with a centre coordinate value $\boldsymbol{w}_c$ from the origin of $\mathbb{R}^4$. Ellipse W defines the spectral band weights that preserve our variance from linearly mapping subject to our linear equality constraints. We know that the positive-definite quadratic form defines the equation of an ellipse, with the positive-definite matrix defining the ratio and orientation of its major and minor semi-axis. Using Eq. (6.23) with Eq. (6.18) we can project $\boldsymbol{\Sigma}$ onto L_o and this provides us with the positive-definite matrix $\boldsymbol{\Sigma}_N$ that is needed to describe W

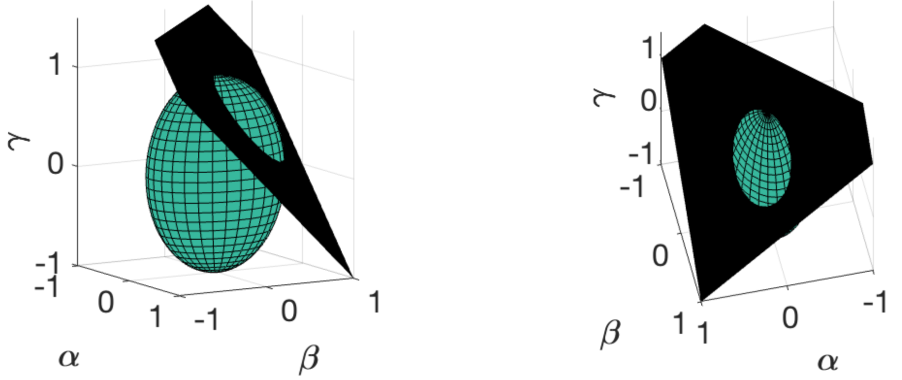
$$X_o = \{\boldsymbol{w}^T \Sigma \boldsymbol{w} \in L_o : \boldsymbol{g}^T \boldsymbol{N}^T \boldsymbol{\Sigma} \boldsymbol{N} \boldsymbol{g} = \boldsymbol{g}^T \boldsymbol{\Sigma}_N \boldsymbol{g} = K_p\} \subseteq \mathbb{R}^4, \tag{6.24}$$

where $\boldsymbol{\Sigma}_N = \boldsymbol{N}^T \boldsymbol{\Sigma} \boldsymbol{N}$ is the positive-definite matrix that we need to describe our ellipse W and the set $X_o = L_o \cap S$ (see Figure. 6.2(e)).

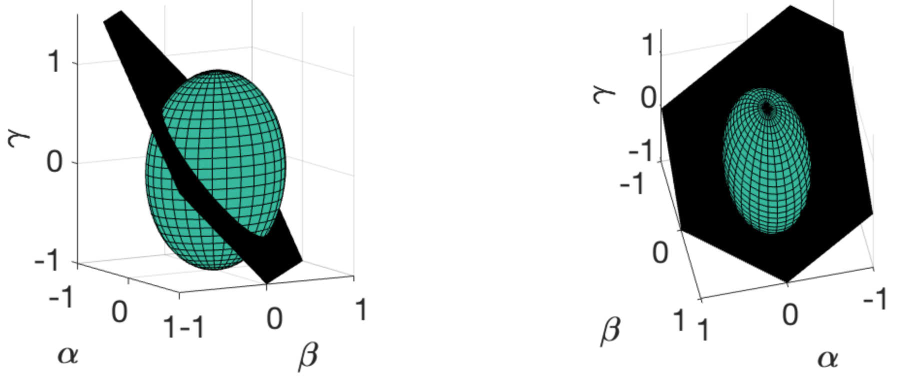
This basis $\boldsymbol{N}$ is computed as a subspace that passes through the origin of $\mathbb{R}^4$ which means that Eq. (6.24) describes the ellipse X_o that possesses axis magnitudes that are larger (although the ratio and orientation are correct) than the one we seek to describe W (an ellipse formed from a plane intersecting an origin centred ellipsoid will be largest when the plane intersects the origin, see Figure. 6.2(e)-(f)). From the construction of our constraints E and B, L will not naturally pass through the origin, which means that the ellipse W that we seek is smaller than that computed by Eq. (6.24) (see Figure. 6.2(d) and (f)). We therefore must adjust the value of the variance we wish to preserve (K_p) so that the major and minor semi-axis of X_o are reduced to the same major and minor semi-axis as W. We shall now describe how we alter K_p so that it produces the ellipse that has the same major and minor semi-axis as W.



(a) S; the ellipsoid describing the variance we wish to preserve (b) $L = E \cap B$; plane produced by intersection of the equality constraints



(c) $W = L \cap S$; the intersection of the equality constraints with the ellipsoid S (d) W looking down the normal of L; this is the desired ellipse subject to our equality constraints



(e) $X_o = L_o \cap\ S$; translation of our equality constraints to the origin (f) X_o looking down the normal of L_o; translation causes an increase in size of our ellipse

Figure 6.2: Geometric representation of our sets and their intersections. Subfigures (d) and (f) are the same as (c) and (e) respectively. The different perspective is provided to aid understanding.

6.4.1 Minimum Variance and the Hyper-Ellipse Centre

Subject to our linear equality constraints, the centre of our ellipse $\boldsymbol{w}_c$ can be described by the minimum variance ($\boldsymbol{w}_c^T \boldsymbol{\Sigma} \boldsymbol{w}_c = K_c < K_p$) in $\mathbb{R}^4$. At this point L is tangent to a smaller ellipsoid

$$S_c = \{\boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} = K_c\} \subseteq \mathbb{R}^4, \tag{6.25}$$

where the normal vector on S_c at $\boldsymbol{w}_c$ is parallel to the vector that is normal to L.

To find this point in $\mathbb{R}^4$ that defines the lowest variance subject to these equality constraints we must solve the following optimisation

$$\min_{\boldsymbol{w}} \quad \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \quad \text{s.t.} \quad \boldsymbol{w} \in L, \tag{6.26}$$

which we can do by again using the Lagrangian Multiplier method

$$L(\boldsymbol{w}, \boldsymbol{\lambda}) = \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} - \lambda_1 \left(\boldsymbol{w}^T \boldsymbol{u} - 1\right) - \lambda_2 \left(\boldsymbol{w}^T \boldsymbol{\mu} - \boldsymbol{u}^T \frac{\boldsymbol{\mu}}{4}\right). \tag{6.27}$$

As before, we differentiate Eq. (6.27) with respect to each variable

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{w}^T} &= 2\boldsymbol{\Sigma} \boldsymbol{w} - \lambda_1 \boldsymbol{u} - \lambda_2 \boldsymbol{\mu} = 0, \\ \frac{\partial L}{\partial \lambda_1} &= \boldsymbol{w}^T \boldsymbol{u} - 1 = 0, \\ \frac{\partial L}{\partial \lambda_2} &= \boldsymbol{w}^T \boldsymbol{\mu} - \boldsymbol{u}^T \frac{\boldsymbol{\mu}}{4} = 0 \end{aligned} \tag{6.28}$$

and solving a system of linear equations provides the general solution

$$\begin{bmatrix} \boldsymbol{w}_c \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} 2\boldsymbol{\Sigma} & -\boldsymbol{A}^T \\ \boldsymbol{A} & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \boldsymbol{k} \end{bmatrix}, \tag{6.29}$$

where $\boldsymbol{w}_c$ is the vector that defines the solution for minimum variance in $\mathbb{R}^4$, $\boldsymbol{\lambda}$ the vector containing the Lagrangian Multipliers and lastly $\boldsymbol{k}$ and $\boldsymbol{A}$ a vector and matrix that contain the constants and coefficients of our linear constraints respectively.

Now that we know $\boldsymbol{w}_c$, we can compute K_c using the quadratic form in Eq. (6.26).

In translating L to the origin, the ellipsoid defined by a variance of K_c goes to zero. In order to preserve the proportions of the major and minor semi-axis of S and W, we subtract K_c from K_p to form ellipsoid S^{-} (see Figure. 6.3(c))

$$S^{-} = \{\boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T \Sigma \boldsymbol{w} = K\} \subseteq \mathbb{R}^4, \tag{6.30}$$

where $K = K_p - K_c$.

The intersection of S^{-} with L_o gives us the ellipse we seek (W) at the origin (W_o) of our vector space, $W_o = S^{-} \cap L_o$, see Figure. 6.3(c)-(d). So,

$$W_o = \{\boldsymbol{g} \in L : \boldsymbol{g}^T \boldsymbol{\Sigma}_N \boldsymbol{g} = K\} \subseteq \mathbb{R}^4. \tag{6.31}$$

Now we have the ellipse we seek at the origin, we can use SVD

$$\boldsymbol{\Sigma}_N = \boldsymbol{U}_N \boldsymbol{S}_N \boldsymbol{U}_N^T. \tag{6.32}$$

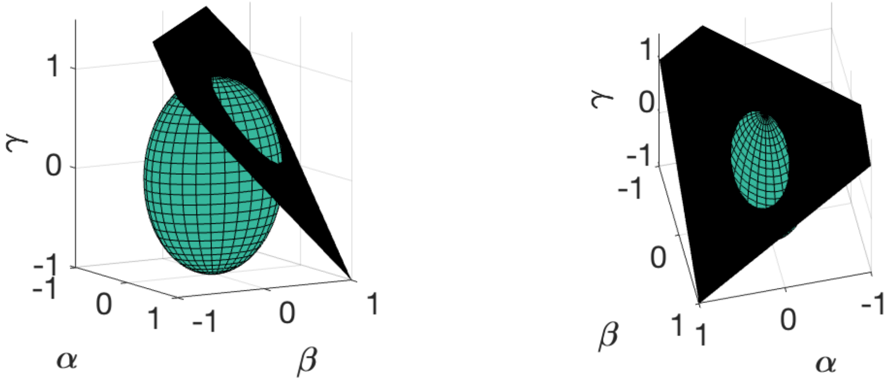
Following this we are now in a position to follow the steps in Section (6.3) to obtain the major and minor semi-axis of W_o

$$\boldsymbol{M}_N = \sqrt{K} \boldsymbol{U}_N \sqrt{\boldsymbol{S}}_N^{-1}. \tag{6.33}$$

Algorithm. 4 summarises these steps to compute the major and minor semi-axis of the ellipse/plane intersection that we have just described in this subsection.

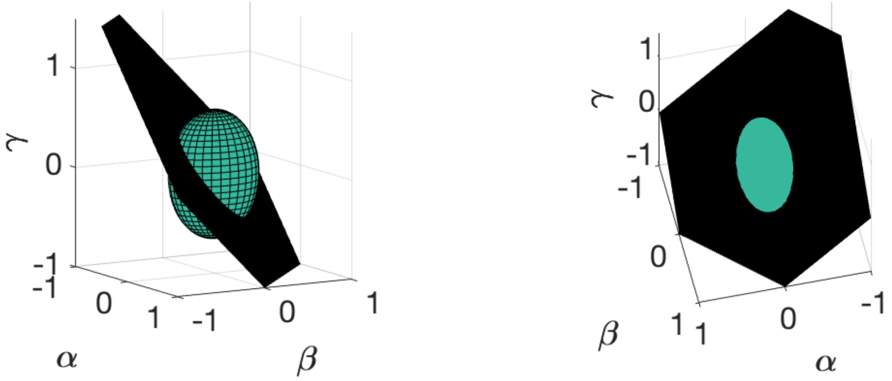
Algorithm 4 Ellipsoid-Plane Intersection.

```
procedure NULLELLIPSE
    Input: Image covariance matrix $\boldsymbol{\Sigma}$, null-space basis $\boldsymbol{N}$ of equality constraints.
    Output: Matrix that contains the semi-minor/major axis $\boldsymbol{M}_N$ of the ellipse generated from the plane/ellipsoid intersection.
    $\boldsymbol{w}_c \leftarrow \min_{\boldsymbol{w}} \quad \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}$        ▷ Solve using Lagrangian Multiplier method Eq. (6.29)
    $K_c \leftarrow \boldsymbol{w}_c^T \boldsymbol{\Sigma} \boldsymbol{w}_c$
    K $\leftarrow K_p - K_c$        ▷ $K_p$ is the variance we wish to preserve
    $\boldsymbol{\Sigma}_N \leftarrow \boldsymbol{N}^T \boldsymbol{\Sigma} \boldsymbol{N}$
    $\boldsymbol{U}_N \boldsymbol{S}_N \boldsymbol{U}_N^T \leftarrow \boldsymbol{\Sigma}_N$        ▷ singular value decomposition
    $\boldsymbol{M}_N \leftarrow \sqrt{K} \boldsymbol{U}_N \sqrt{\boldsymbol{S}}_N^{-1}$
    Return $\boldsymbol{M}_N, \boldsymbol{U}_N, \boldsymbol{S}_N, \boldsymbol{\Sigma}_N, K$
```



(a) $W = L \cap S$; the intersection of the equality constraints with the ellipsoid S. Ellipse centre is at $\boldsymbol{w}_c \in \mathbb{R}^3$

(b) W looking down the normal of L; this is the desired ellipse subject to our equality constraints. $\boldsymbol{w}_c$ is the centre of the ellipse



(c) $W_o = S^{-} \cap L_o$; S reduced by change in variance produced from moving $\boldsymbol{w}_c$ to origin

(d) W_o looking down the normal of L_o; our desired ellipse centred at the origin of our vector space

Figure 6.3: Geometric representation of our sets and their intersections. Subfigure (b) is the same as (a). The different perspective is provided to aid understanding.

6.4.2 Ellipse and Line Segment Intersections

In Chapter 3 we described the intersection of half-spaces to provide a convex set (that we call Q in this chapter) that defined band weights for linearly mapping a colour image to its greyscale. This problem was solved for a 1-dimensional space L and our convex set was simply a line segment. Solving the same problem with four bands means that L is now 2-dimensional; the edges of our convex set are now formed of multiple line segments. The intersection of these edges with the ellipse W (in addition to the elements of $W \in Q$) will define the band weights that will preserve our variance K_p for our linear mapping.

To calculate our intersection of the line segments that make up the edges of Q_o (Q_o is our convex set Q that has been translated to the origin by $\boldsymbol{w}_c$), we must formulate W_o into the general form of an ellipse (so we don't have any bilinear terms in our equation, see Section 6.3). Let's begin by defining our rotation function

$$R : \boldsymbol{g} \in L : L_o \to D, \boldsymbol{g} \mapsto \boldsymbol{U}_N \boldsymbol{g} \tag{6.34}$$

and its inverse

$$R^{-1} : \boldsymbol{d} \in D : D \to L_o, \boldsymbol{d} \mapsto \boldsymbol{U}_N^{-1} \boldsymbol{d}, \tag{6.35}$$

where D is our rotated subspace L and $\boldsymbol{d}$ represent all of the rotated $\boldsymbol{g}$ vectors. Applying this rotation to our ellipse W_o

$$R(W_o) = \{\boldsymbol{d} \in D : \boldsymbol{d}^T \boldsymbol{S}_N \boldsymbol{d} = K\} \subseteq \mathbb{R}^4. \tag{6.36}$$

Let us express the equations of our line segments as

$$R(Q_o) = \{\rho \in \mathbb{R} : \boldsymbol{d} = \rho \boldsymbol{d}_y + (1 - \rho)\boldsymbol{d}_x, 0 \leq \rho \leq 1\} \tag{6.37}$$

where ρ is the mixing coefficient in $\mathbb{R}$ and $\boldsymbol{d}_x$ and $\boldsymbol{d}_y$ are the rotated vertices of Q_o. Substituting these line segments into our quadratic form that describes our ellipse $R(W_o)$ results in a quadratic equation of the form (Finlayson and Matheson, 2012),

$$x\rho^2 + y\rho + z = 0, \tag{6.38}$$

where

$$\begin{aligned} x &= \left(\boldsymbol{d}_x^T \boldsymbol{S}_N \boldsymbol{d}_x\right) - 2\left(\boldsymbol{d}_x^T \boldsymbol{S}_N \boldsymbol{d}_y\right) + \left(\boldsymbol{d}_y^T \boldsymbol{S}_N \boldsymbol{d}_y\right), \\ y &= \left(\boldsymbol{d}_x^T \boldsymbol{S}_N \boldsymbol{d}_y\right) - \left(\boldsymbol{d}_x^T \boldsymbol{S}_N \boldsymbol{d}_x\right), \\ z &= \left(\boldsymbol{d}_x^T \boldsymbol{S}_N \boldsymbol{d}_x\right) - K, \end{aligned} \tag{6.39}$$

which can be solved for ρ using the quadratic equation

$$\rho = \frac{-y \pm \sqrt{y^2 - 4xz}}{2x}. \tag{6.40}$$

The value $y^2 - 4xz$ is called the determinant and the number of real solutions to the equation depends on whether the determinant is positive; the line will intersect twice, negative; the line and ellipse do not intersect or zero; the line and ellipse are tangent (Stephens, 2017, Weisstein, 2019a, Weisstein, 2019b). Filtering out the lines with determinants less than zero is shown by lines 8 to 10 of Algorithm. 5 and only allows ρ to be calculated for lines that will intercept the ellipse.

Substituting the remaining coefficients from Eq. (6.39) into the quadratic equation (Eq. (6.40)) provides the values of ρ where the line defined by the vectors $\boldsymbol{d}_x$ and $\boldsymbol{d}_y$ intersect with the ellipse defined by Eq. (6.36). As we are only interested in the points where the line segment intersect (as only the segment make up the boundaries of our convex set and not the entire line itself) we dismiss lines that correspond with ρ values less than 0 or greater than 1 (lines 11 to 12 of Algorithm. 5).

Substituting our computed values of ρ from Eq. (6.40) into the equation of our line segment Eq. (6.37) gives us our ellipse/line-segment intersection, $\boldsymbol{d}^x$ (line 13 of Algorithm. 5).

Algorithm 5 Ellipse Line-Segment Intersections.

```
1: procedure ELLIPSELINEX
2:     Input: $\boldsymbol{S}, K$ and radially ordered vertices of convex set $\mathbf{Q}$
3:     Output: line segment intersections with ellipse, $\boldsymbol{d}^{\mathbf{x}}$
4:     Insert a copy of the first column of $\mathbf{Q}$ as (n+1) column of $\mathbf{Q}$
5:     $\mathbf{Q}_i$, $\mathbf{Q}_{i+1} \leftarrow$ vertices of line segment
6:     for i from 1 to length of $\mathbf{Q} - 1$ do
7:         $x_i, y_i, z_i \leftarrow$ Solve coefficients of contrast form, Eq. (6.39)
8:         if $y_i^2 - 4x_i z_i < 0$ then
9:             $\emptyset \leftarrow x_i, y_i, z_i$
10:        $\rho_i \leftarrow \left(-y_i \pm \sqrt{y_i^2 - 4x_i z_i}\right)/2x_i$
11:        if $0 > \rho_i > 1$ then
12:            $\emptyset \leftarrow \rho_i$
13:        $\boldsymbol{d}_i^{\mathbf{x}} \leftarrow \mathbf{Q}_{i+1} + (1 - \rho_i)\mathbf{Q}_i$
14:    Return $\boldsymbol{d}^{\mathbf{x}}$
```

6.4.3 Enhanced Correlation

In Section 6.2.1 we described how to calculate the band weights ($\boldsymbol{w}_{cor}$) that produces a greyscale that has maximum correlation with $\boldsymbol{P}$. A problem arises in that $\boldsymbol{w}_{cor}$ is very unlikely to lie on our ellipse W. As we wish to prioritise preserving the image variance, we must instead find the point on W that would produce a correlation that is as close to the one produced with the band weights defined by $\boldsymbol{w}_{cor}$. Like variance, correlation can be described by a quadratic form which, in two dimensions, describes an ellipse. Recall that a correlation matrix is just a normalised covariance matrix and will possess the same eigenvectors and eigenvalues. In other words, with regard to variance and correlation, the metric space is anisotropic, the incremental change in correlation/variance depends on the vector direction through the space. In this section we shall show that we can simply use the Euclidean metric from the origin to find the shortest distance to the ellipse W. To apply the Euclidean magnitude as our metric of distance we have transform the vector space such that it is isotropic, the incremental change in variance/correlation is independent of direction and we can do this by transforming the covariance matrix into the identity matrix $\boldsymbol{I}$ (Golub and Loan, 1996, Finlayson and Matheson, 2012).

With reference to the pseudocode shown by Algorithm. 6, we use the mappings described in Section 6.4 to map $\boldsymbol{w}_{cor}$ to $\boldsymbol{d}_{cor}$ (line 6 of Algorithm. 6) and here the positive-definite matrix that describes the elliptical space is described by the eigenvalues ($\boldsymbol{S}_N$) of $\boldsymbol{\Sigma}_N$. In this elliptical space, $\boldsymbol{d}_{cor}$ will define a point on the ellipse $K_{cor} = \boldsymbol{d}_{cor}^T \boldsymbol{S}_N \boldsymbol{d}_{cor}$ (line 7 of Algorithm. 6)

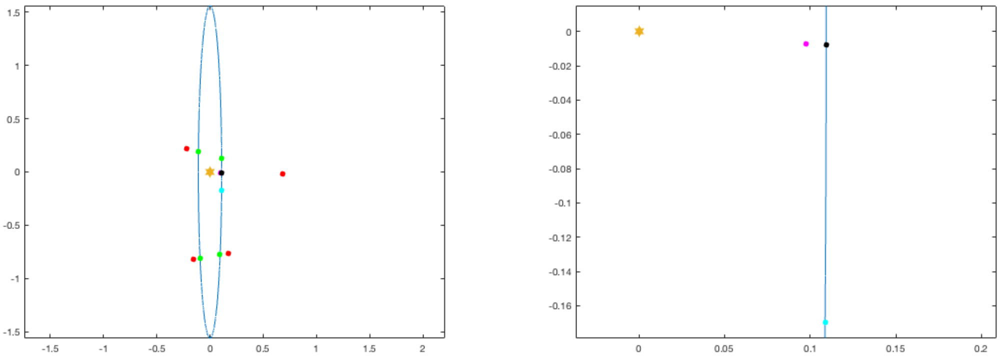
$$W_o^{cor} = \{\boldsymbol{d} \in D : \boldsymbol{d}^T \boldsymbol{S}_N \boldsymbol{d} = K_{cor}\} \subseteq \mathbb{R}^4, \tag{6.41}$$

which will sit on the circle where our metric space is isotropic

$$O_o^{cor} = \{\boldsymbol{d} \in D : \boldsymbol{d}^T \sqrt{\boldsymbol{S}}_N^T I \sqrt{\boldsymbol{S}}_N \boldsymbol{d} = K_{cor}\} \subseteq \mathbb{R}^4. \tag{6.42}$$

Transforming our elliptical space to a spherical one also means that the ellipse that describes our variance K becomes

$$O_o = \{\boldsymbol{d} \in D : \boldsymbol{d}^T \sqrt{\boldsymbol{S}}_N^T I \sqrt{\boldsymbol{S}}_N \boldsymbol{d} = K\} \subseteq \mathbb{R}^4. \tag{6.43}$$



(a) Ellipse that defines our variance s.t. our equality constraints

(b) Close-up to show resolution between $\boldsymbol{d}_{cor}$(magenta) and $\boldsymbol{d}^{\sim}_{cor}$ (black)

Figure 6.4: Example of the geometry of our ellipse $R(W_o)$ in D. The coloured markers are as follows: Brown = ellipse centre, red = vertices of $R(Q_o)$, green = $\boldsymbol{d}^x$, cyan = $\boldsymbol{d}_I = \boldsymbol{u}/4 \rightarrow \boldsymbol{d}$, magenta = $\boldsymbol{d}_{cor}$, black = $\boldsymbol{d}^{\sim}_{cor}$.

The Euclidean magnitude from $\sqrt{\boldsymbol{S}}_N\boldsymbol{d}_{cor}$ to the circle that describes our transformed variance O_o is a straight line through the origin. So by modulating $\sqrt{\boldsymbol{S}}\boldsymbol{d}_{cor}$ so that it has a magnitude K, we can find the intersection with O_o to be

$$\sqrt{\frac{K}{K_{cor}}}\sqrt{\boldsymbol{S}}_N\boldsymbol{d}_{cor}. \tag{6.44}$$

Transforming the vector space (and the vector described by Eq. (6.46)) back to an ellipsoid to give us the vector

$$\boldsymbol{d}^{\sim}_{cor} = \sqrt{\boldsymbol{S}}_N^{-1}\sqrt{\frac{K}{K_{cor}}}\sqrt{\boldsymbol{S}}_N\boldsymbol{d}_{cor}, \tag{6.45}$$

which cancels down to (line 7 of Algorithm. 6)

$$\boldsymbol{d}^{\sim}_{cor} = \sqrt{\frac{K}{K_{cor}}}\boldsymbol{d}_{cor}, \tag{6.46}$$

which is a vector whose inner product is scaled by the ratio of K to K_{cor}. In other words, the vector $\boldsymbol{d}_{cor}$ has been linearly scaled such the vector $\boldsymbol{d}^{\sim}_{cor}$ lies on our ellipse W_o, see Figure. 6.4.

All that is left to do is to check if this vector lies within the elliptical line segment defined by the vectors $\boldsymbol{d}^x$. If it does, then we have our solution. If it does not, then we select the $\boldsymbol{d}^x$ closest to it. Either way, we map our vector back to $\mathbb{R}^4$ to obtain our final band weights. So, inverting

Algorithm 6 Enhanced Correlation.

```
procedure ENHCOR
    Input: Multiband image $\boldsymbol{M}$, panchromatic image $\boldsymbol{P}$, basis $\boldsymbol{N}$
    Output: vector that defines band weights to preserve variance and enhanced correlation $\boldsymbol{d}^{\sim}_{cor}$
    Compute $\boldsymbol{w}_{cor}$                                        ▷ Eq. (6.13)
    $\boldsymbol{U}_N, \boldsymbol{S}_N, \boldsymbol{w}_c, K \leftarrow \mathrm{nullEllipse}\,(\boldsymbol{\Sigma}, \boldsymbol{N})$     ▷ Algorithm. 4
    Map $\boldsymbol{w}_{cor}$ to $\boldsymbol{d}_{cor} \leftarrow \boldsymbol{U}_N^T * ((\boldsymbol{w}_{cor} - \boldsymbol{w}_c)^T * \boldsymbol{N})^T$
    $K_{cor} \leftarrow \boldsymbol{d}_{cor}^T \boldsymbol{S}_N \boldsymbol{d}_{cor}$
    Compute $\boldsymbol{d}^{\sim}_{cor}$                                  ▷ Eq. (6.46)
    Return $\boldsymbol{d}^{\sim}_{cor}$
```

our rotation; $R^{-1}(\boldsymbol{d}^x) : \boldsymbol{d}^x \rightarrow \boldsymbol{g}^x$, mapping back to $\mathbb{R}^4$; $F'(\boldsymbol{g}^x) : \boldsymbol{g}^x \rightarrow \boldsymbol{w}_o^x$ and translating back by $\boldsymbol{w}_c$ to obtain our line segment ellipse intersections, $\boldsymbol{w}^x$; $\boldsymbol{w} = \boldsymbol{w}_o^x + \boldsymbol{w}_c$ in our original vector space V.

We include Algorithm. 7 as our final pseudocode that brings together all of our work in this chapter for producing our approximation of the panchromatic image. To aid the reader we shall review the process.

Lines 5 to 11 computes the basis of the intersection of our equality constraints, where we check for degeneracy (line 5) of our brightness preserving constraint (discussed in the next subsection).

Lines 12 and 13 defines and computes the intersection of our half-spaces (Section 3.2.2) in the basis of these equality constraints; the polytope formed from this step defines all the possible band weightings available to produce a linearly mapped and visually meaningful greyscale subject to our constraints.

Line 14 calls Algorithm. 4; our procedure for computing the intersection of our linear and quadratic equality constraints. This involved solving for the intersection of hyper-planes with a hyper-ellipse; a solution we presented in Section 6.4.1.

Lines 15 to 17 takes the vertices that we computed from line 13, maps them back to our original vector space (Eq. (6.35)), translates them by $\boldsymbol{w}_c$ and maps them back (Eq. (6.21)) to the basis defined by the intersection of our equality constraints (line 16). These translated vertices are finally rotated (Subsection 6.4.2 Eq. (6.34)) by the eigenvector matrix $\boldsymbol{U}_N$.

Line 18 calls Algorithm. 5; our procedure for computing where our convex polytope formed from the vertices in $\boldsymbol{d}$ intersects the our ellipse defined on line 14. We discussed this in Subsection. 6.4.2.

Line 19 calls Algorithm. 6; our procedure for computing the band weights for producing our greyscale $\boldsymbol{J}$ subject to all of our constraints that has a greater correlation than $\boldsymbol{I}$ with $\boldsymbol{P}$.

Lines 21 to 24 ensure that these band weights from line 19 produce a greyscale with no pixel clipping. If it does, the closest point (using the Euclidean metric) where our convex polytope intersects our variance preserving constraint is used.

Lines 25 to 28 reverses all of our mappings (Eq. (6.23)), translations and rotations (Eq. (6.35)) to bring our final set of band weights (from lines 21 to 24) back to our original vector space.

Algorithm 7 Contrast and Brightness Preserving Panchromatic Approximation.

```
 1: procedure CONTRASTBRIGHTNESSPRESERVEPAN
 2:     Input: Spectral bands as an array of pixel inequalities $P$, covariance matrix $\boldsymbol{\Sigma}$ and the
        panchromatic image $\boldsymbol{P}$
 3:     Output: Greyscale image $\boldsymbol{J}$ and the vector that contains the band weights $\boldsymbol{w}$
 4:     $\boldsymbol{I}^T \leftarrow [1/4,\ 1/4,\ 1/4,\ 1/4]$                              ▷ Interior point
 5:     if $\boldsymbol{\mu}_R = \boldsymbol{\mu}_G = \boldsymbol{\mu}_B = \boldsymbol{\mu}_{NIR}$ then                        ▷ Degenerate case
 6:         $\boldsymbol{A} \leftarrow \left(\left[\boldsymbol{u},\ \boldsymbol{\mu},\ \boldsymbol{\psi}\right]^T\right)$
 7:         $\boldsymbol{k} \leftarrow \left[1, \boldsymbol{\mu}_I, \psi_I\right]^T$
 8:     else
 9:         $\boldsymbol{A} \leftarrow \left(\left[\boldsymbol{u},\ \boldsymbol{\mu}\right]^T\right)$
10:         $\boldsymbol{k} \leftarrow \left[1, \boldsymbol{\mu}_I\right]^T$
11:     $\boldsymbol{N} \leftarrow \mathrm{Null}(\boldsymbol{A})$
12:     $\Pi \leftarrow [P\boldsymbol{N},\ (P\boldsymbol{I} + b)]$                                ▷ $\Pi = P \cap L$
13:     $\boldsymbol{v} \leftarrow \mathrm{intersectHalfspaces}(\Pi)$                          ▷ Algorithm. 1
14:     $\boldsymbol{S}_N, \boldsymbol{U}_N, \boldsymbol{w}_c, K \leftarrow \mathrm{nullEllipse}(\boldsymbol{\Sigma}, \boldsymbol{N})$                ▷ Algorithm. 4
15:     $\boldsymbol{w}^T \leftarrow \left(\boldsymbol{N}^T\mathbf{v} + \boldsymbol{I}^T\right) - \boldsymbol{w}_c$
16:     $\boldsymbol{g} \leftarrow \boldsymbol{N}^T\boldsymbol{w}$
17:     $\boldsymbol{d} \leftarrow \boldsymbol{U}_N\boldsymbol{g}$
18:     $\boldsymbol{d}^x \leftarrow \mathrm{EllipseLineX}(\boldsymbol{S}_N, K, \boldsymbol{d})$                      ▷ Algorithm. 5
19:     $\boldsymbol{d}^{\sim}_{cor} \leftarrow \mathrm{enhCor}(\boldsymbol{M}, \boldsymbol{P})$                            ▷ Algorithm. 6
20:     insert $\boldsymbol{d}^{\sim}_{cor}$ as last vector in $\boldsymbol{d}^x$
21:     if $\boldsymbol{d}^x$ has 5 vectors then
22:         $\boldsymbol{d}^x \leftarrow$ vector associated with the median angle from the radially sorted vectors in $\boldsymbol{d}^x$
        that coincide with the same sign as $\boldsymbol{d}^{\sim}_{cor}$
23:     else
24:         $\boldsymbol{d}^x \leftarrow$ vector associated with the median angle from the radially sorted vectors in $\boldsymbol{d}^x$
25:     $\boldsymbol{g}^x \leftarrow \boldsymbol{U}_N^{-1}\boldsymbol{d}^x$
26:     $\boldsymbol{w}_o^x \leftarrow \boldsymbol{N}\boldsymbol{g}^x$
27:     $\boldsymbol{w} \leftarrow \boldsymbol{w}_o^x + \boldsymbol{w}_c$
28:     Construct $\boldsymbol{J}$ using the band weights contained in $\boldsymbol{w}$
29:     Return $\boldsymbol{w}$, $\boldsymbol{J}$
```

6.4.4 Degenerate Case

Unlike the colour-to-greyscale case, degeneracy in the equality constraints is not a simple matter of increasing the dimensionality of our convex set and finding the correct vertex. This is due to our solution being reliant on finding the intersection of a line segment with an ellipse. For example, if we were to add just one extra dimension to L then it would possess 3 basis vectors in $\mathbb{R}^4$ which increases the dimensionality of our convex set; the line segments that define the edges of Q become convex faces, our ellipse W becomes an ellipsoid and their intersection creates ellipses and not points. The option available to us in these instances would be to add another equality constraint to reduce L back to 2 dimensions. One such constraint would be to add one that allows for us to control the range (ψ) of our $\boldsymbol{J}$ mapping. For example, if we decided to preserve the range of $\boldsymbol{I}$ then

$$R = \left\{ \boldsymbol{w} \in \mathbb{R}^4 : \boldsymbol{w}^T \boldsymbol{\psi} = \boldsymbol{u}^T \frac{\boldsymbol{\psi}}{4} \right\} \subseteq \mathbb{R}^4, \tag{6.47}$$

where

$$\boldsymbol{\psi}^T = [\psi_R \quad \psi_G \quad \psi_B \quad \psi_{NIR}]. \tag{6.48}$$

There are two points worth noting with this constraint: 1) it can be applied even when we are not dealing with the degenerate case, and this would reduce the basis of L down to a single vector in $\mathbb{R}^4$. Consequently we would then be looking at where this single line segment in L intersects W to find our solution. 2) This constraint is only viable provided that at least one band in $\boldsymbol{M}$ has a range different to 1. Otherwise it would be degenerate with the energy preserving constraint.

6.5 Results and Discussion

We tested our mean and variance preserving multispectral-to-greyscale conversion using five images collected from three different satellites. Images of Rio de Janeiro; 512×512 (Figure. 6.5 (a) and (b)) and Tripoli; 512×512 (Figure. 6.7 (a) and (b)) were taken with the World-View 3 (WV-3), Stockholm; 512×512 (Figure. 6.9 (a) and (b)) and Washington DC 2048×2048 (Figure. 6.11

(a) and (b)) the World-View 2 (WV-2) and lastly, Toulouse; 1024×1024 (Figure. 6.13 (a) and (b)) with the Pléiades. In these figures, subfigure (c) represents our greyscale mapping $\boldsymbol{J}$. We have not included the intensity mapping $\boldsymbol{I}$ as there is no significantly noticeable difference with $\boldsymbol{J}$. The differences between the two (d) and the difference in their edge maps (f) are better portrayed using the ratio method described in the subfigure captions.

Table. 6.1 shows the results for our greyscale mappings $\boldsymbol{J}$ which, with respect to $\boldsymbol{I}$, possesses a greater than or equal correlation with $\boldsymbol{P}$. For the Toulouse image, we see a maximum increase in correlation of 0.15%. Interestingly for this particular image there is no contribution of the green band in producing $\boldsymbol{J}$; a consequence of prioritising preserving the global statistics over maximising correlation. Additionally it is worth noting that for this image only $\boldsymbol{P}$ was synthetically generated by the authors by averaging the green and the red bands followed by a series of filtering steps. For the remaining images the increase in correlation is smaller and is especially noticeable for the Tripoli image which sees no significant (at 4 decimal place accuracy) difference from $\boldsymbol{I}$, a result that is expected given that the band weights for this image are the closest to those of $\boldsymbol{I}$ than the remaining four. This is portrayed with the difference map in Figure. 6.7(d) where the image is dominated with grey values of 0.5, which is to say that $\boldsymbol{J} \approx \boldsymbol{I}$. For the remaining difference maps, despite there being only a small deviation in the correlations at 3 to 4 decimal places, more variation from 0.5 can be seen as random noise scatter or image edge definition.

To show the differences between the edge maps produced using $\boldsymbol{P} - \boldsymbol{J}$ and $\boldsymbol{P} - \boldsymbol{I}$ we have

Table 6.1: Multispectral band weights for $\boldsymbol{J}$ and correlation data with the panchromatic image across the five datasets.

Satellite	Image	Weights for $\boldsymbol{J}$				Correlation with $\boldsymbol{P}$	
		α	β	γ	ϵ	$\boldsymbol{I}$	$\boldsymbol{J}$
WV-3	Rio de Janeiro	0.2748	0.2790	0.2359	0.2102	0.9517	0.9519
	Tripoli	0.2580	0.2363	0.2555	0.2501	0.8527	0.8527
WV-2	Stockholm	0.3090	0.1152	0.3171	0.2588	0.8709	0.8710
	Washington DC	0.3115	0.1861	0.2436	0.2528	0.7978	0.7979
Pléiades	Toulouse	0.3113	0	0.4226	0.2661	0.8737	0.8750

calculated images using the ratio

$$\frac{\boldsymbol{P}-\boldsymbol{J}}{(\boldsymbol{P}-\boldsymbol{I})+(\boldsymbol{P}-\boldsymbol{J})}=\frac{\boldsymbol{P}-\boldsymbol{J}}{2\boldsymbol{P}-\boldsymbol{I}-\boldsymbol{J}}. \tag{6.49}$$

Computing the ratio of the edge maps in this way means that when $\boldsymbol{I}=\boldsymbol{J}$ then the calculation reduces to produce a greyscale pixel value of 0.5. Instances when $\boldsymbol{P}=\boldsymbol{J}$ will result in a black pixel and conversely a white pixel is produced when $\boldsymbol{P}=\boldsymbol{I}$. Subfigure (f) in Figures. (6.5), (6.7), (6.9), (6.11) and (6.13) show these differences in the edge maps where the largest difference can be seen for the Rio de Janeiro and Toulouse datasets.

Subfigures (a) and (b) in Figures. (6.6), (6.8), (6.10), (6.12) and (6.14) show the geometry of our solutions and the image histograms for $\boldsymbol{I}$ (subfigure (c)), $\boldsymbol{J}$ (subfigure (d)) and $\boldsymbol{P}$ (subfigure (e)). The geometry of our solution shows us looking down the normal of D where we see ellipse $R(W_o)$, the vertices $\boldsymbol{d}$ (red) of the convex set $R(Q)$, the intersection of this convex set with $R(W_o)$ (green), the vector of highest correlation subject to our equality constraints $\boldsymbol{d}_{cor}$ (magenta) and its projection to W_o from the origin $\boldsymbol{w}_c$, $\boldsymbol{d}^{\sim}_{cor}$ (black) and lastly the transformation of the intensity mapping vector to D (cyan).

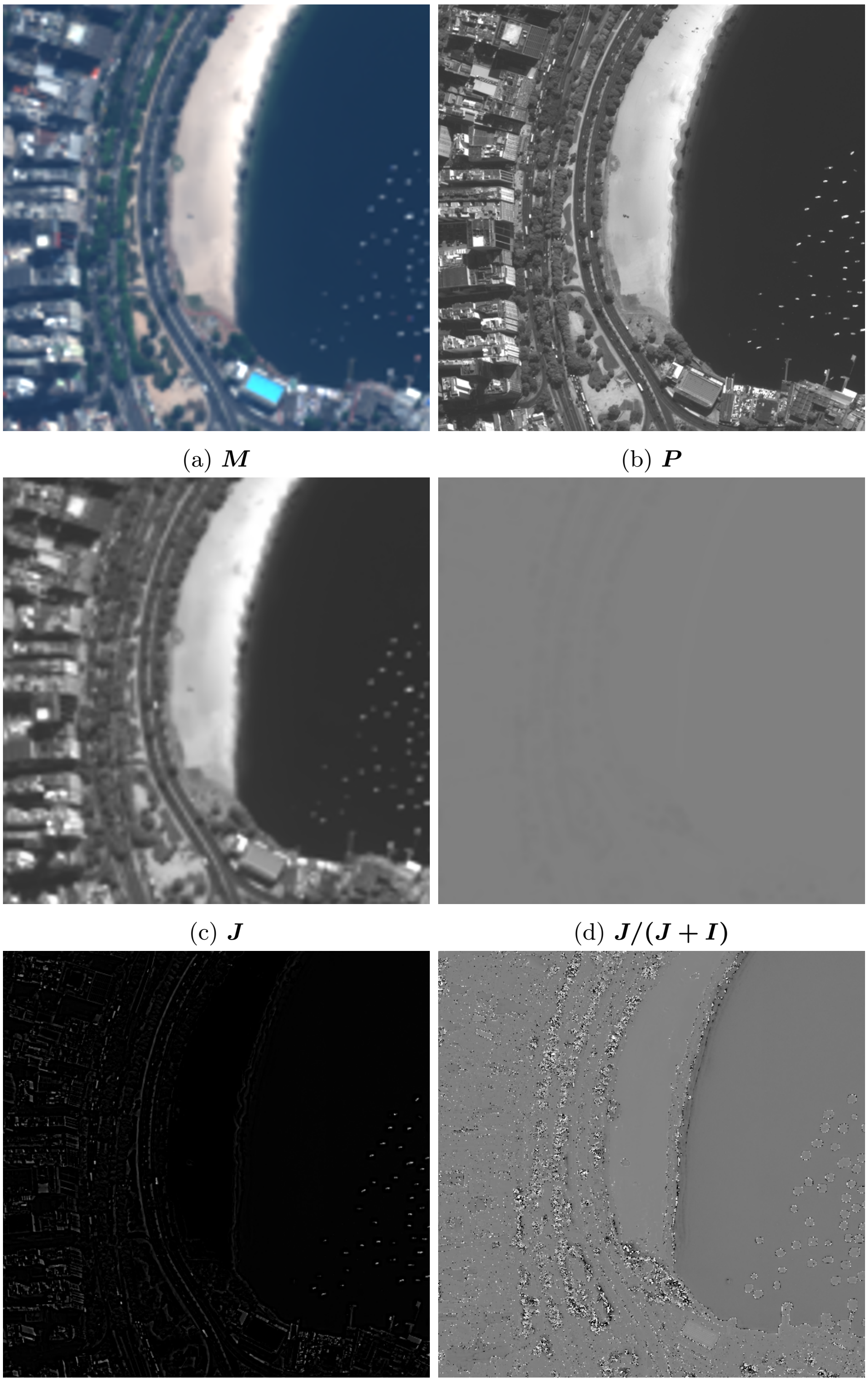
As we have previously said, the images where $\boldsymbol{d}^{\sim}_{cor}$ exists outside the elliptical line segments (defined by the intersection of W_o with the convex set Q) will produce band weights for $\boldsymbol{J}$ that will cause image clipping. In this instance we dismiss it and take the elliptical line segment vertex that is closest to $\boldsymbol{d}^{\sim}_{cor}$. This occurs for two images; Tripoli and Toulouse (Figures. 6.8(b) and 6.14(b)). For these images we select the bottom and top line segment intersections (green) respectively. Coincidently these images also possess the respective lowest and largest correlation differences with $\boldsymbol{P}$.

The remaining three images produce not only $\boldsymbol{d}^{\sim}_{cor}$ within Q but also $\boldsymbol{d}_{cor}$. This means that should we wish to release our constraint on variance, we could produce a normalised image $\boldsymbol{J}$ with an even greater correlation with $\boldsymbol{P}$ while preserving the mean of $\boldsymbol{M}$.

With respect to $\boldsymbol{I}$, the Tripoli image shows very little difference in band weights and correlation (see Table. 6.1) with $\boldsymbol{P}$ and this is reflected in subfigure (b) where $\boldsymbol{d}^{\sim}_{cor}$ approximates $\boldsymbol{d}_I$, with a difference of 0.02 in their y vector components. This is in contrast to the Toulouse image, which has a far larger difference of approximately 0.5.

Given the very similar measures of correlation for $\boldsymbol{J}$ and $\boldsymbol{I}$ to $\boldsymbol{P}$ it is expected that their histograms look nearly identical to each other. The highest correlation to $\boldsymbol{P}$ was for the image Rio de Janeiro (0.9519) and this can be visually seen in how close the histograms of $\boldsymbol{J}$, $\boldsymbol{I}$ and $\boldsymbol{P}$ all appear for this dataset. Disappointingly $\boldsymbol{P}$ contains pixel values greater than 1; a side effect from the image pre-processing statistical matching procedure. This puts our solution at an immediate handicap as our intersection of half-spaces computation restricts our convex set Q to only allow for band weights to take values between 0 and 1. This problem also restricts our solution for the Stockholm image in addition to the peak at around a pixel value of 0.2, where $\boldsymbol{I}$ and $\boldsymbol{J}$ have approximately 10,000 pixels missing (which accounts for approximately 4% of the image) from this bin.

The remaining image histograms show greater dispersion from $\boldsymbol{P}$. The Washington DC dataset gives the lowest correlation for $\boldsymbol{I}$ and $\boldsymbol{J}$ to $\boldsymbol{P}$ which we attribute to the presence of additional pixel values in the 0.2 pixel bin that are missing from $\boldsymbol{P}$. Further to this disparity, $\boldsymbol{P}$ appears to possess a significant amount of pixels at around the 0 and 0.82 values that are missing from $\boldsymbol{J}$.



(a) $\boldsymbol{M}$

(b) $\boldsymbol{P}$

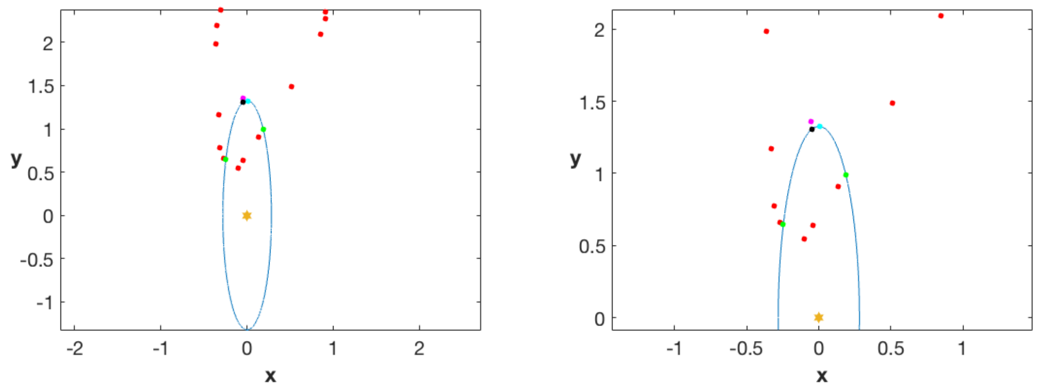
(c) $\boldsymbol{J}$

(d) $\boldsymbol{J}/(\boldsymbol{J}+\boldsymbol{I})$

(e) $\boldsymbol{P}-\boldsymbol{J}$

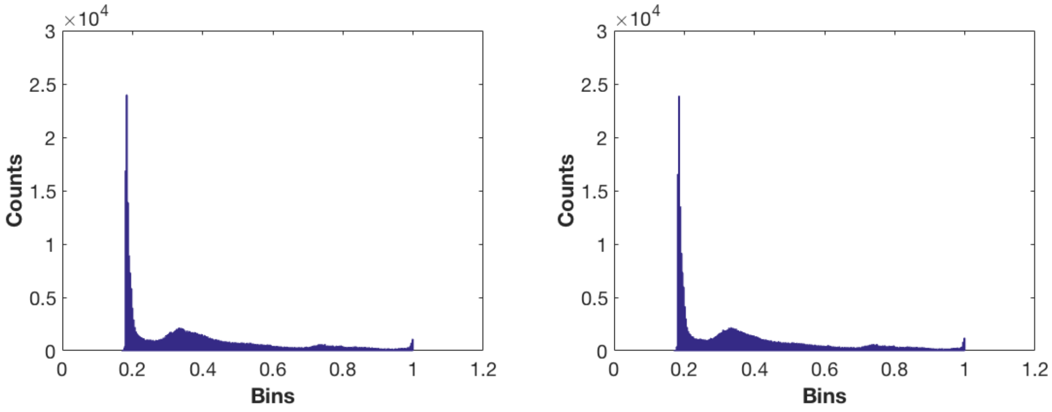
(f) $(\boldsymbol{P}-\boldsymbol{J})/(\mathbf{2}\boldsymbol{P}-\boldsymbol{I}-\boldsymbol{J})$

Figure 6.5: Rio de Janeiro captured from the WV-3 satellite (a) RGB, (b) panchromatic, (c) our linear mapping $\boldsymbol{J}$ that preserves variance and mean of $\boldsymbol{I}$, (d) differences between $\boldsymbol{J}$ and $\boldsymbol{I}$, (e) edge map produced from $\boldsymbol{P}-\boldsymbol{J}$ and (f) difference between the edge maps $\boldsymbol{P}-\boldsymbol{J}$ and $\mathbf{P}-\mathbf{I}$.



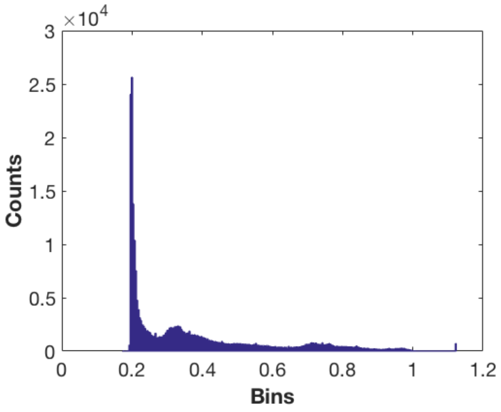
(a) Geometry of our solution

(b) Close-up geometry of our solution



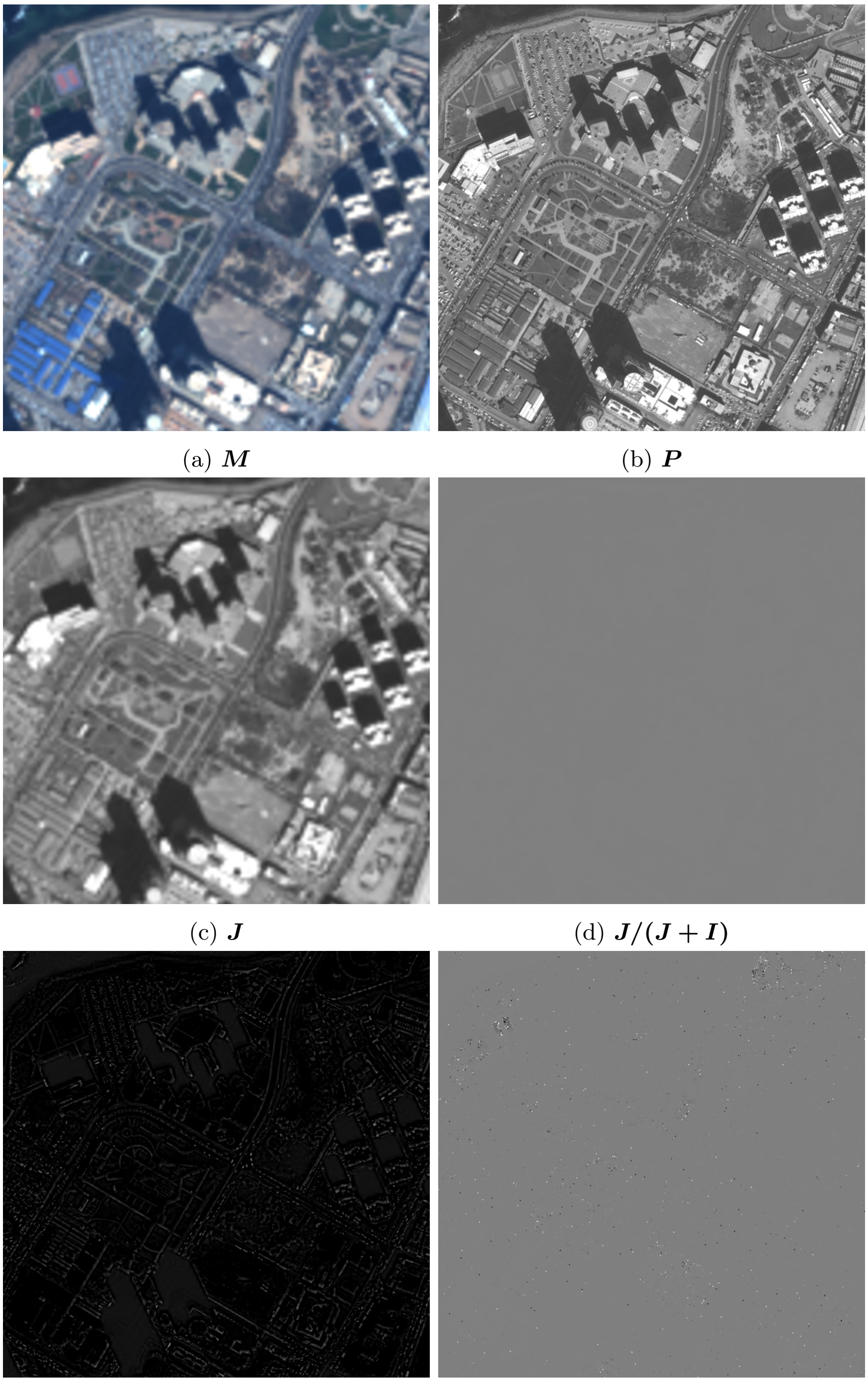
(c) Histogram of $\boldsymbol{I}$

(d) Histogram of $\boldsymbol{J}$



(e) Histogram of $\boldsymbol{P}$

Figure 6.6: Rio de Janeiro: (a)-(b) Geometry of our ellipse $R(W_o)$ in D. The coloured markers are as follows: Brown = ellipse centre, red = vertices of $R(Q_o)$, green = $\boldsymbol{d}^x$, cyan = $\boldsymbol{u}/4 \rightarrow \boldsymbol{d}$, magenta = $\boldsymbol{d}_{cor}$, black = $\boldsymbol{d}_{cor}^{\sim}$, (c) $\boldsymbol{I}$, (d) $\boldsymbol{J}$ and (e) $\boldsymbol{P}$ histograms.



(a) $\boldsymbol{M}$

(b) $\boldsymbol{P}$

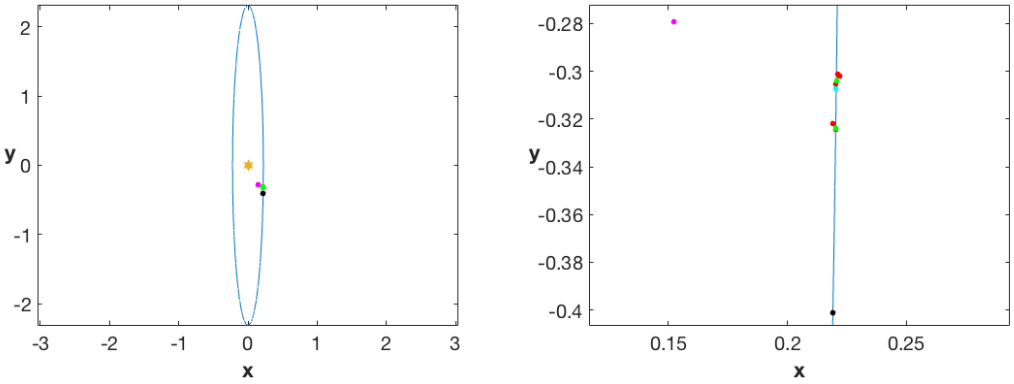
(c) $\boldsymbol{J}$

(d) $\boldsymbol{J}/(\boldsymbol{J}+\boldsymbol{I})$

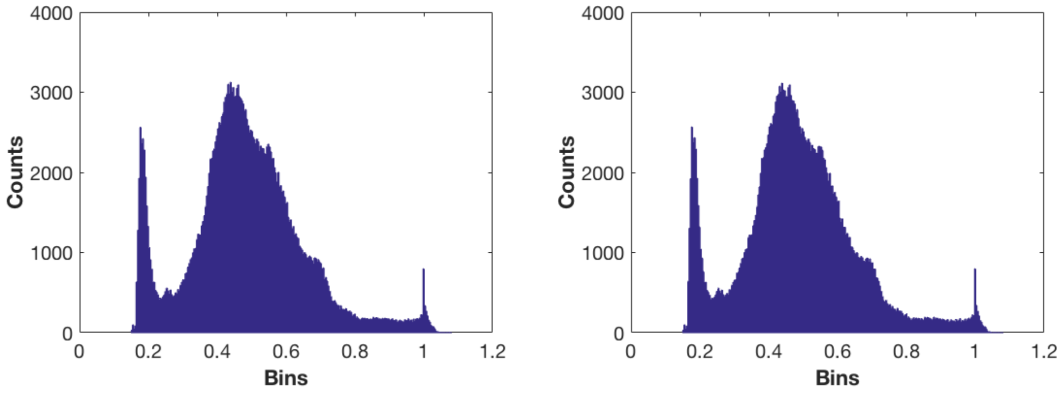
(e) $\boldsymbol{P}-\boldsymbol{J}$

(f) $(\boldsymbol{P}-\boldsymbol{J})/(2\boldsymbol{P}-\boldsymbol{I}-\boldsymbol{J})$

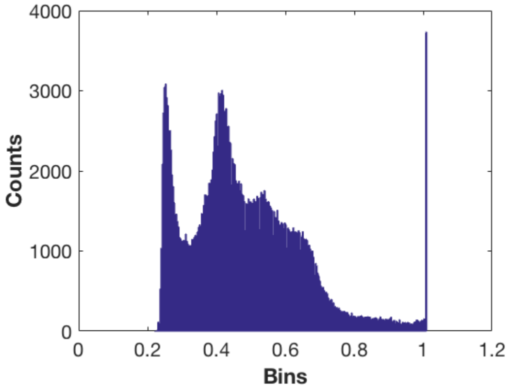
Figure 6.7: Tripoli captured from the WV-3 satellite (a) RGB, (b) panchromatic, (c) our linear mapping $\boldsymbol{J}$ that preserves variance and mean of $\boldsymbol{I}$, (d) differences between $\boldsymbol{J}$ and $\boldsymbol{I}$, (e) edge map produced from $\boldsymbol{P}-\boldsymbol{J}$ and (f) difference between the edge maps $\boldsymbol{P}-\boldsymbol{J}$ and $\mathbf{P}-\mathbf{I}$.



(a) Geometry of our solution (b) Geometry of our solution: Enlarged

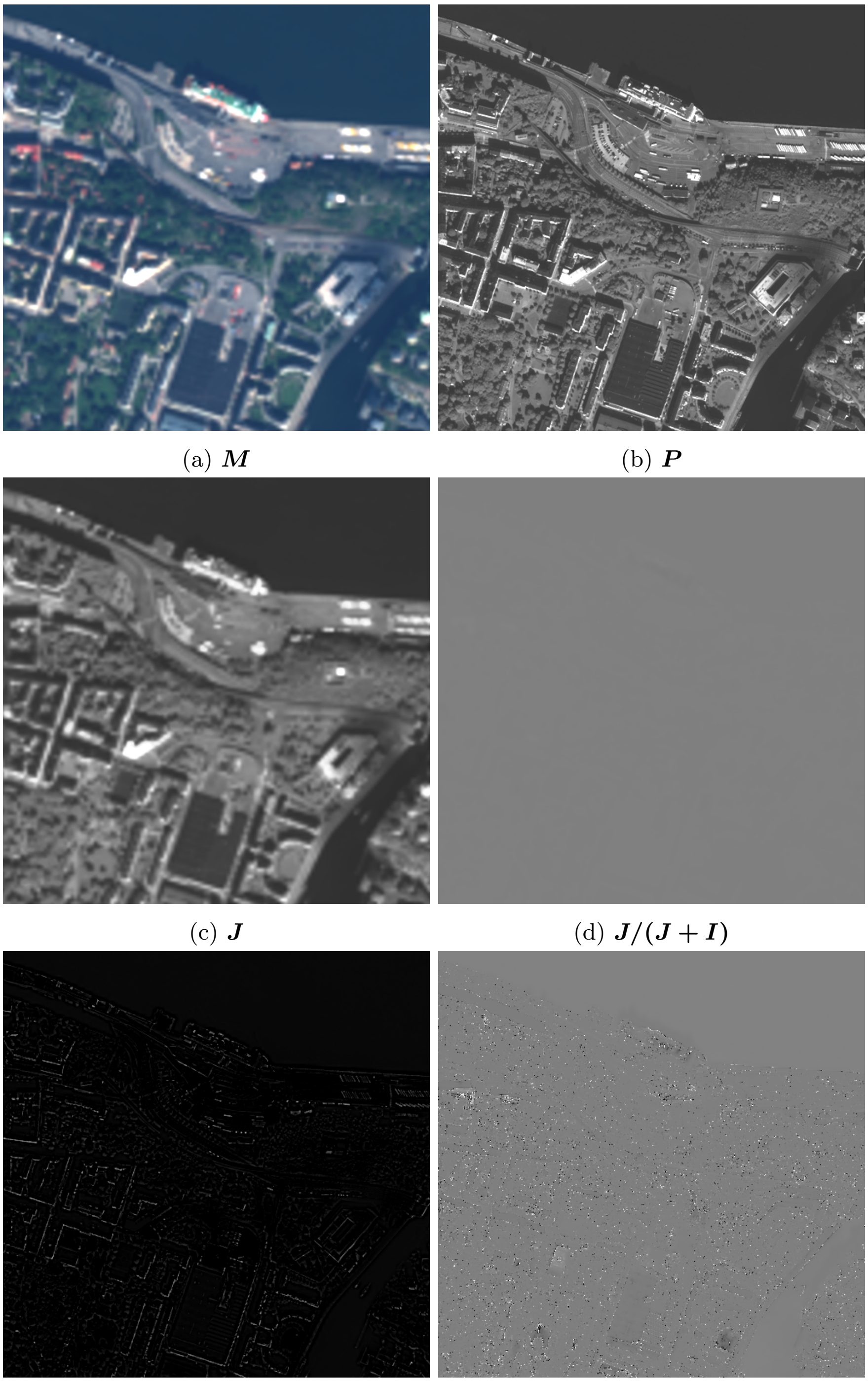


(c) Histogram of $\boldsymbol{I}$ (d) Histogram of $\boldsymbol{J}$



(e) Histogram of $\boldsymbol{P}$

Figure 6.8: Tripoli: (a)-(b) Geometry of our ellipse $R(W_o)$ in D. The coloured markers are as follows: Brown = ellipse centre, red = vertices of $R(Q_o)$, green = $\boldsymbol{d}^x$, cyan = $\boldsymbol{u}/4 \rightarrow \boldsymbol{d}$, magenta = $\boldsymbol{d}_{cor}$, black = $\boldsymbol{d}^{\sim}_{cor}$, (c) $\boldsymbol{I}$, (d) $\boldsymbol{J}$ and (e) $\boldsymbol{P}$ histograms.

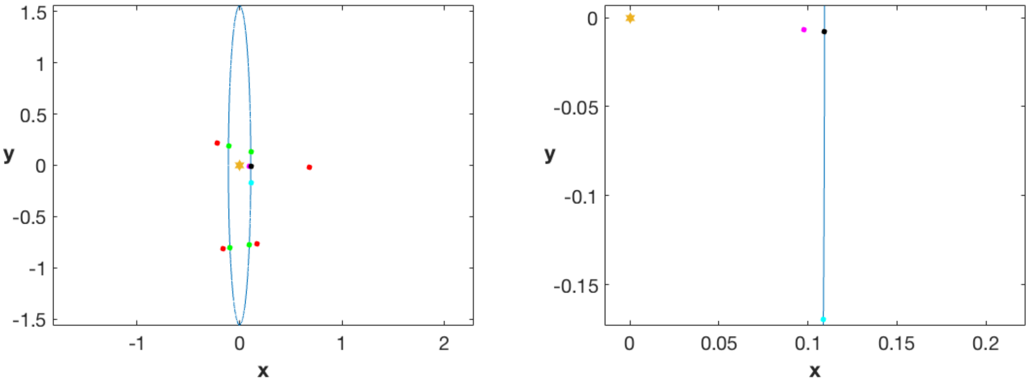


(a) $\boldsymbol{M}$ (b) $\boldsymbol{P}$

(c) $\boldsymbol{J}$ (d) $\boldsymbol{J/(J+I)}$

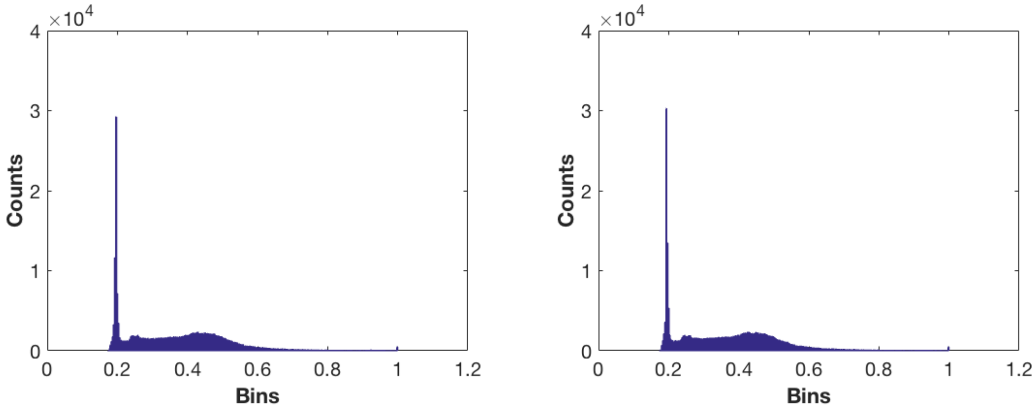
(e) $\boldsymbol{P-J}$ (f) $\boldsymbol{(P-J)/(2P-I-J)}$

Figure 6.9: Stockholm captured from the WV-2 satellite (a) RGB, (b) panchromatic, (c) our linear mapping $\boldsymbol{J}$ that preserves variance and mean of $\boldsymbol{I}$, (d) differences between $\boldsymbol{J}$ and $\boldsymbol{I}$, (e) edge map produced from $\boldsymbol{P-J}$ and (f) difference between the edge maps $\boldsymbol{P-J}$ and $\mathbf{P}-\mathbf{I}$.



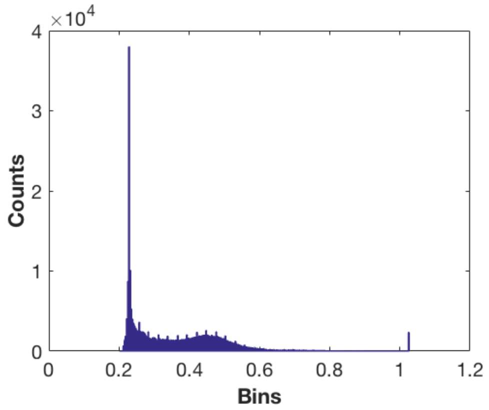
(a) Geometry of our solution

(b) Geometry of our solution: Enlarged



(c) Histogram of $\boldsymbol{I}$

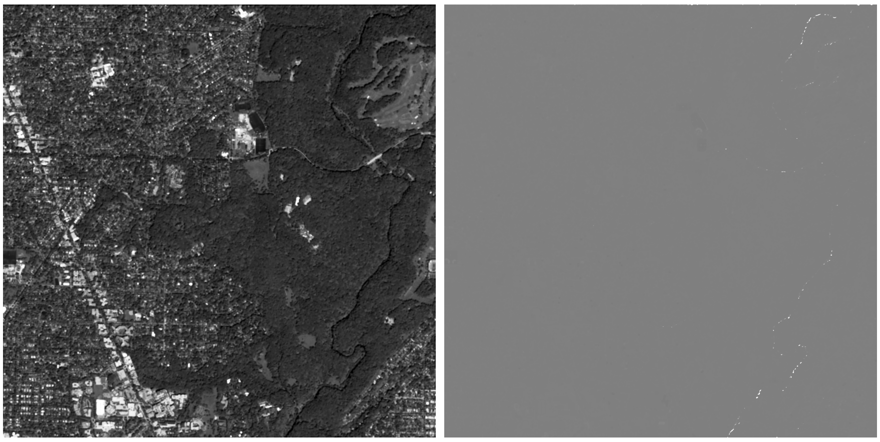
(d) Histogram of $\boldsymbol{J}$



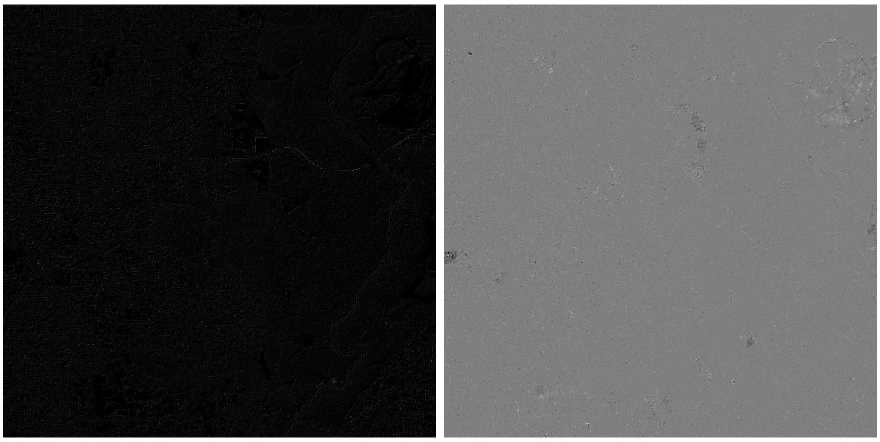
(e) Histogram of $\boldsymbol{P}$

Figure 6.10: Stockholm: (a) Geometry of our ellipse $R(W_o)$ in D. The coloured markers are as follows: Brown = ellipse centre, red = vertices of $R(Q_o)$, green = $\boldsymbol{d}^x$, cyan = $\boldsymbol{u}/4 \to \boldsymbol{d}$, magenta = $\boldsymbol{d}_{cor}$, black = $\boldsymbol{d}^{\sim}_{cor}$, (c) $\boldsymbol{I}$, (d) $\boldsymbol{J}$ and (e) $\boldsymbol{P}$ histograms.

(a) $\boldsymbol{M}$ (b) $\boldsymbol{P}$

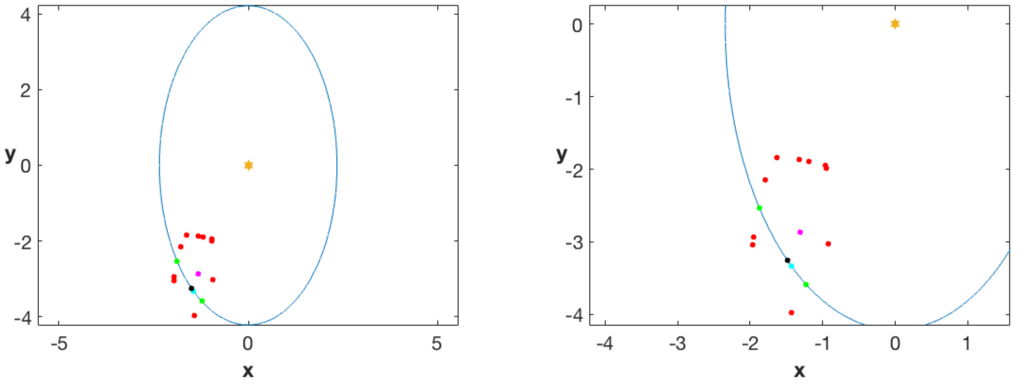


(c) $\boldsymbol{J}$ (d) $\boldsymbol{J/(J+I)}$



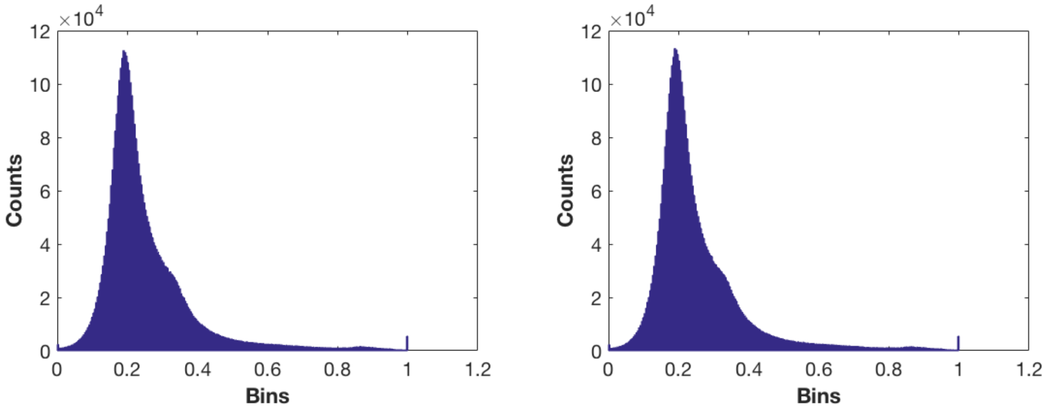
(e) $\boldsymbol{P-J}$ (f) $\boldsymbol{(P-J)/(2P-I-J)}$

Figure 6.11: Washington DC captured from the WV-2 satellite (a) RGB, (b) panchromatic, (c) our linear mapping $\boldsymbol{J}$ that preserves variance and mean of $\boldsymbol{I}$ (d) differences between $\boldsymbol{J}$ and $\boldsymbol{I}$, (e) edge map produced from $\boldsymbol{P-J}$ and (f) difference between the edge maps $\boldsymbol{P-J}$ and $\mathbf{P}-\mathbf{I}$.



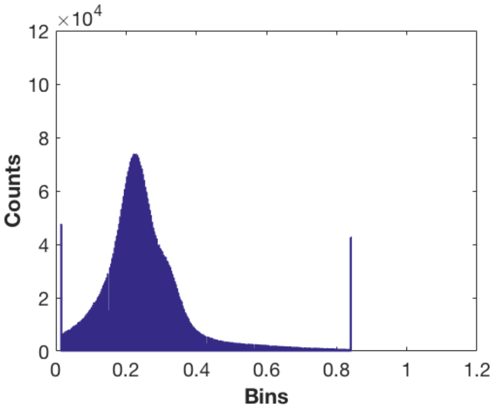
(a) Geometry of our solution

(b) Geometry of our solution: Enlarged



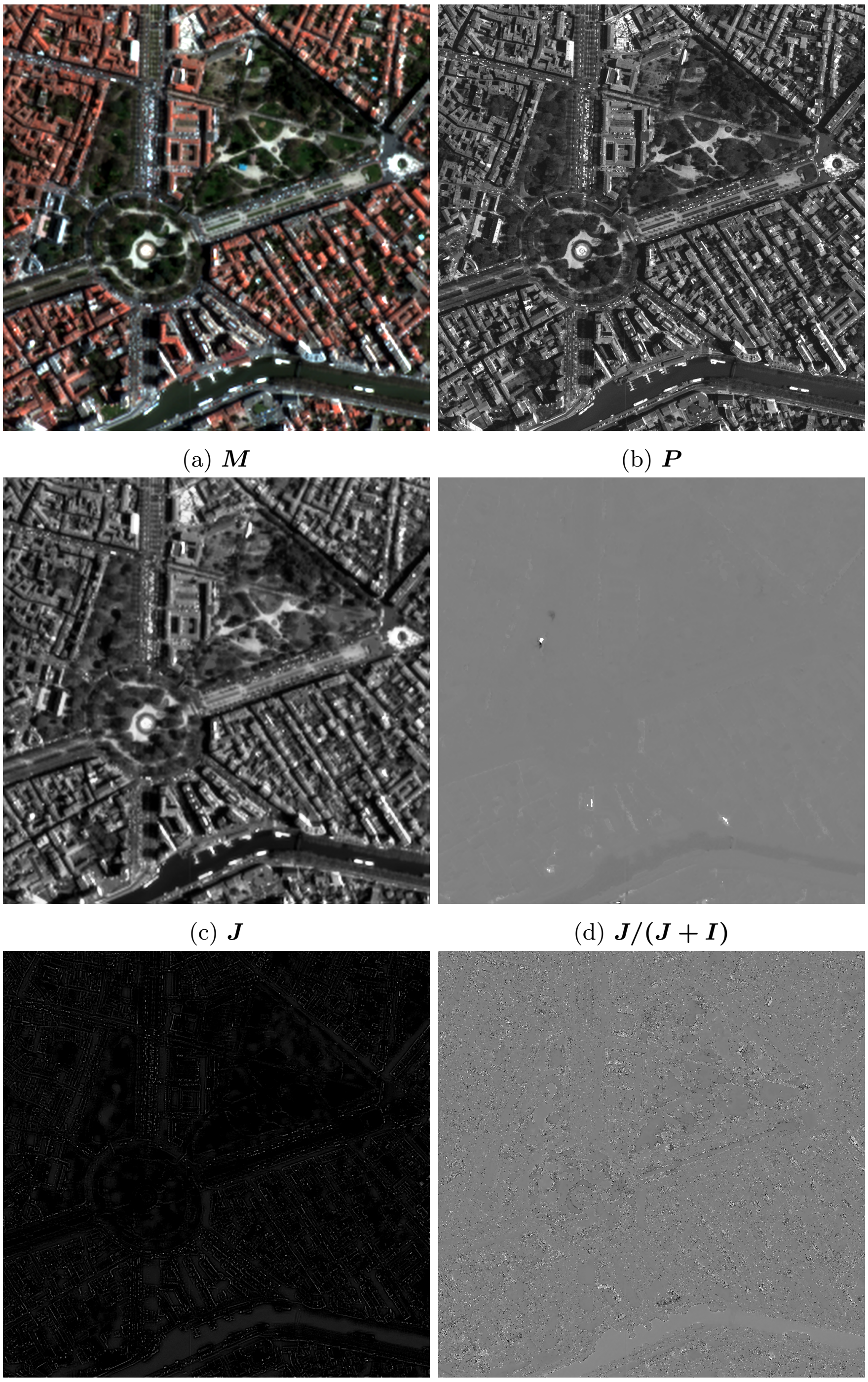
(c) Histogram of $\boldsymbol{I}$

(d) Histogram of $\boldsymbol{J}$



(e) Histogram of $\boldsymbol{P}$

Figure 6.12: Washington DC: (a) Geometry of our ellipse $R(W_o)$ in D. The coloured markers are as follows: Brown = ellipse centre, red = vertices of $R(Q_o)$, green = $\boldsymbol{d}^x$, cyan = $\boldsymbol{u}/4 \rightarrow \boldsymbol{d}$, magenta = $\boldsymbol{d}_{cor}$, black = $\boldsymbol{d}^{\sim}_{cor}$, (c) $\boldsymbol{I}$, (d) $\boldsymbol{J}$ and (e) $\boldsymbol{P}$ histograms.

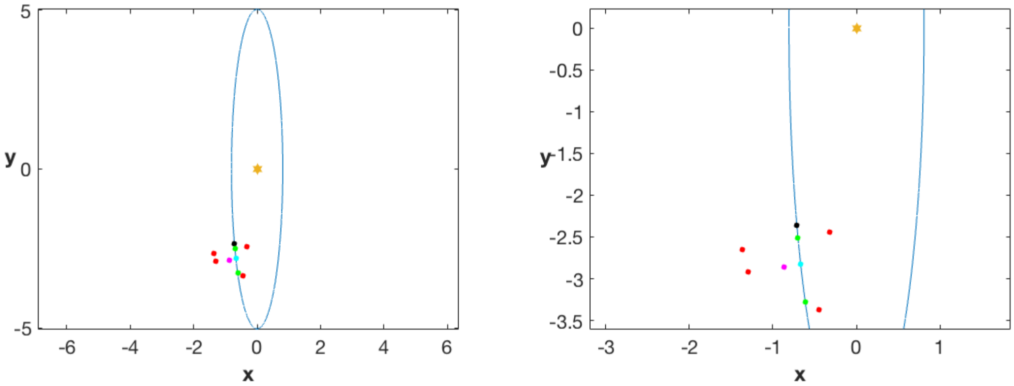


(a) $\boldsymbol{M}$ (b) $\boldsymbol{P}$

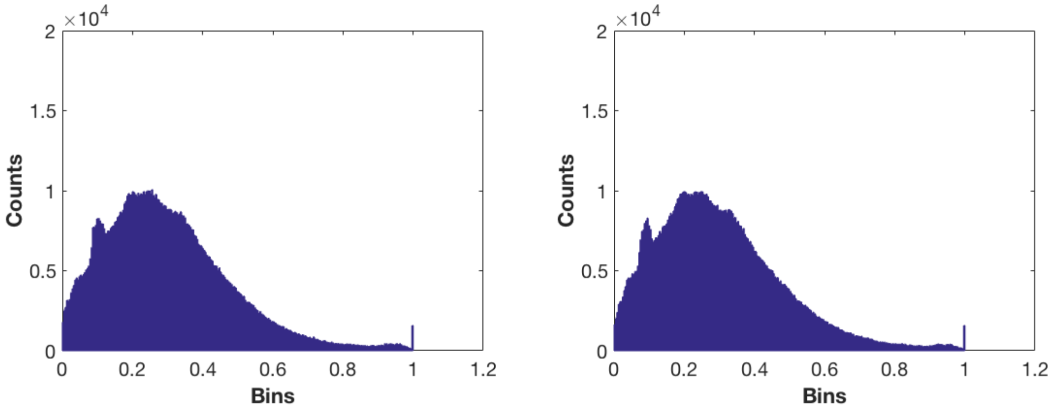
(c) $\boldsymbol{J}$ (d) $\boldsymbol{J}/(\boldsymbol{J}+\boldsymbol{I})$

(e) $\boldsymbol{P}-\boldsymbol{J}$ (f) $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$

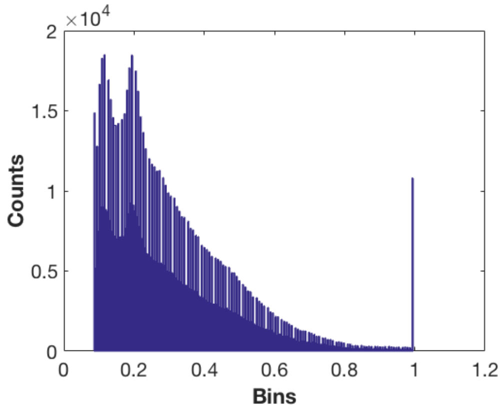
Figure 6.13: Toulouse captured from the Pléiades satellite (a) RGB, (b) panchromatic, (c) our linear mapping $\boldsymbol{J}$ that preserves variance and mean of $\boldsymbol{I}$, (d) differences between $\boldsymbol{J}$ and $\boldsymbol{I}$, (e) edge map produced from $\boldsymbol{P}-\boldsymbol{J}$ and (f) difference between the edge maps $\boldsymbol{P}-\boldsymbol{J}$ and $\mathbf{P}-\mathbf{I}$.



(a) Geometry of our solution (b) Geometry of our solution: Enlarged



(c) Histogram of $\boldsymbol{I}$ (d) Histogram of $\boldsymbol{J}$



(e) Histogram of $\boldsymbol{P}$

Figure 6.14: Toulouse: (a) Geometry of our ellipse $R(W_o)$ in D. The coloured markers are as follows: Brown = ellipse centre, red = vertices of $R(Q_o)$, green = $\boldsymbol{d}^x$, cyan = $\boldsymbol{u}/4 \rightarrow \boldsymbol{d}$, magenta = $\boldsymbol{d}_{cor}$, black = $\boldsymbol{d}_{cor}^{\sim}$, (c) $\boldsymbol{I}$, (d) $\boldsymbol{J}$ and (e) $\boldsymbol{P}$ histograms.

6.6 Complexity Analysis

In this section we shall discuss the time complexity of our algorithm. Our algorithm relies on three functions: singular value decomposition (SVD), convex hull and matrix inversion. The complexity is determined by the method that is used to carry out the function, for example, in Chapter 3 we mentioned that if the mergesort was used instead of quicksort then the complexity would change from $O(n^2)$ to $O(n\ log\ n)$. Let us assume the worst case for the SVD and matrix inversion and that both have a complexity of $O(d^3)$ (Pan and Chen (1999), Tian et al. (2014)), where d is the dimensionality of the square matrices.

In our algorithm SVD can be used to compute the null spaces of our two 3-dimensional hyper-planes in $\mathbb{R}^4$ by finding the eigenvectors associated with the zero eigenvalue. Additionally we carry out SVD on a 2×2 matrix when we compute the eigenvalues of our positive-definite matrices. For our constrained minima (using the Lagrangian Multiplier method) optimisation we invert a 6×6 matrix and this is the upper limit for our worst case; a complexity of $O(6^3)$. The time it takes to compute the convex hull in the plane depends on the number of different pixels (n) in an image, with the fastest method having a complexity of $O(n\ log\ n)$. This means that we would need at least 107 points in the plane for the complexity of the convex hull to equal that of our matrix inversion. The $\boldsymbol{J}$ images that we use in our experiment range have histograms that occupy a minimum of 80% of the 256 available bins between 0 and 1. This means that the range in pixel values for our image is 205 in the worst case which if we assume that no multispectral band pixels are being mapped to the same shade of grey means we have 205 different pixels and 410 points in the plane (2 inequalities per different pixel). As $410 > 216$ we conclude that the convex hull function is the limiting factor for time complexity over that of the functions based on arithmetic operations.

6.7 Conclusion

In this chapter we produced a novel solution to linearly mapping four spectral bands to a single greyscale $\boldsymbol{J}$ such that the mean and variance of the intensity mapping $\boldsymbol{I}$ is preserved. To carry out this task we used multivariate calculus to 1) create a novel solution to the intersection of a hyper-ellipse (or hyper-sphere) with a hyper-plane and 2) produce a method of enhancing the correlation of $\boldsymbol{J}$ with the higher resolution panchromatic image $\boldsymbol{P}$. Both of these solutions can be extendeds to imagery of any number of dimensions d.

To ensure that our image $\boldsymbol{J}$ had brightness values within the dynamic range [0 1] we imposed inequality constraints on our multispectral band ($\boldsymbol{M}$) pixels. Solving these inequalities in an intersection of half-space framework provided for us our convex set that defined band weighting vectors needed to producing a normalised image that preserved the statistical mean of $\boldsymbol{I}$. To preserve image variance the problem became one of finding where this convex set intersected an ellipse. Subject to our linear equality constraints, this problem could be solved in 2 dimensions, (thus limiting the complexity of our method to that of creating convex hulls in the plane) by limiting the number of bands in $\boldsymbol{M}$ to four.

We tested our solution on five datasets from three different satellites and we found that with respect to $\boldsymbol{I}$, $\boldsymbol{J}$ had a greater or equal correlation with $\boldsymbol{P}$, with the highest increase being 0.15%.

In the next chapter we seek to apply our $\boldsymbol{J}$ images to component substitution based pan-sharpening methods. We shall evaluate the performance of doing this using Wald's protocol and standard evaluation metrics.

Chapter 7

Multispectral Panchromatic Sharpening: Evaluation

In the previous chapter we demonstrated that by extending the dimensionality of our vector space from 3 to 4, we could modify our convex model so that it could be used for a brightness preserving multispectral-to-greyscale. We then went on to show how we could further adapt our solution to linearly map a four band image into a normalised greyscale which preserved the variance of the intensity mapping while increasing its correlation with the panchromatic.

It is well known that the effectiveness of component substitution (CS) methods for panchromatic sharpening is limited by the correlation of $\boldsymbol{I}$ with $\boldsymbol{P}$. In this chapter we wish to see if we can improve on these methods by employing our new greater correlated image $\boldsymbol{J}$. Using the the freely available toolbox that accompanies the work of Vivone et al. (2015) as a basis for our experiments, we incorporate our imagery into their implementation and process the results using standard metrics provided by the authors.

This chapter follows a standard experimental methodology. We shall first introduce the process of how pan-sharpened images are evaluated followed by outlining our experimental design. We shall then present the results of the pan-sharpened images and their evaluation scores before finishing with a conclusion.

7.1 Introduction

The aim of panchromatic sharpening algorithms is to inject high-frequency panchromatic information into the lower resolution spectral bands. The ideal outcome from this process is to produce a colour image that would be identical to the one captured had the spatial resolution of the spectral band sensors been the same as the panchromatic sensor. If this high-resolution multispectral image was available, then a sensible post-sharpening evaluation would be to then measure the difference between the two images.

Unfortunately the desired high-resolution image (typically referred to as the *ground-truth*) from which a comparison can be made is not usually available. A solution to this problem would be to degrade the spatial resolution of a copy of the multispectral image followed by upsampling it back to its original size. The original image can then be used as the ground-truth for the pan-sharpened image in accordance with the criteria defined by Wald's protocol (Wald et al., 1997).

7.1.1 Wald's Protocol

Let us summarise Wald's protocol.

1. Our pan-sharpened image $\boldsymbol{H}$, once degraded, should be as close as possible to the original multispectral image $\boldsymbol{M}$.
2. Our pan-sharpened image $\boldsymbol{H}$ should be as close as possible to the ground-truth $\boldsymbol{T}$.
3. Each individual spectral band in our pan-sharpened image $\boldsymbol{H}$ should be as close as possible to the corresponding spectral bands in the ground-truth $\boldsymbol{T}$.

Reducing the resolution is carried out by applying a low-pass filter (LPF) and decimation operator (characterised by a sampling factor equal to resolution ratio between the original panchromatic and multispectral images) on the images. For an accurate representation of the imagery at this reduced resolution, the choice of LPF must me made such that it simulates the modular transfer function (MTF) of the satellite sensor.

To apply Wald's protocol we must use similarity metrics between the pan-sharpened image and the ground-truth. One such metric would be to sum the measure of variation between the vectors at each pixel,

$$\mathrm{SAM}(\boldsymbol{H}_{\mathrm{i}}, \boldsymbol{T}_{\mathrm{i}}) = \arccos \frac{\langle \boldsymbol{H}_{\mathrm{i}}, \boldsymbol{T}_{\mathrm{i}} \rangle}{\|\boldsymbol{H}_{\mathrm{i}}\| \, \|\boldsymbol{T}_{\mathrm{i}}\|} \tag{7.1}$$

which is known as the *Spectral Angle Mapper* (SAM). Here the subscript refers to a pixel, $\langle \cdot, \cdot \rangle$ the inner/scalar product and $\|\cdot\|$ the Euclidean norm. The optimum value is when there is no dispersion for the vectors between each image which would give a sum of 0.

A common metric that measures the similarity of two signals is the *Root-Mean-Square Error* (RMSE),

$$\mathrm{RMSE}(\boldsymbol{H}, \boldsymbol{T}) = \sqrt{E\left[(\boldsymbol{H} - \boldsymbol{T}^2\right]} \tag{7.2}$$

The RMSE is popular as it is used in other metrics like the *Erreur Relative Globale Adimensionnelle de Synthèse* (ERGAS),

$$\mathrm{ERGAS} = \frac{100}{R} \sqrt{\frac{1}{d} \sum_{n=1}^{d} \left(\frac{\mathrm{RMSE}(\boldsymbol{H}_n, \boldsymbol{T}_n)}{\mu_{\boldsymbol{H}_n}} \right)^2} \tag{7.3}$$

which is a scaled average of the band specific RMSE that has been normalised to the bands mean. Another metric of choice that is based on statistical moments is the *Universal Image Quality Index* or *Q-index*. Vivone et al. (2015) have eloquently written it in the form,

$$\mathrm{Q}(\boldsymbol{H}, \boldsymbol{T}) = \frac{\sigma_{\boldsymbol{H},\boldsymbol{T}}}{\sigma_{\boldsymbol{H}} \sigma_{\boldsymbol{T}}} \frac{2\mu_{\boldsymbol{H}} \mu_{\boldsymbol{T}}}{(\mu_{\boldsymbol{H}}^2 + \mu_{\boldsymbol{T}}^2)} \frac{2\sigma_{\boldsymbol{H}} \sigma_{\boldsymbol{T}}}{(\sigma_{\boldsymbol{H}}^2 + \sigma_{\boldsymbol{T}}^2)} \tag{7.4}$$

which, by order of appearance, shows that the metric is based on the correlation, brightness and contrast between the images. The correlation coefficient is the ratio of the covariance to the product of the standard deviations of the images. The standard deviation describes the distribution of data around the mean pixel value. Covariance is the un-normalised measure of the relationship between the bands of the image. Dividing through by the product of the standard deviations of both images normalises the the covariance to lie within the range [-1, 1] and this is what we call the correlation. The second term measures the proximity of the average brightness of the pixels between the compared images and the third describes the similarity of contrast.

Having described our evaluation metrics, we shall now outline our experimental design.

7.2 Experimental Design

The aim of our experiment is to assess whether using a greyscale $\boldsymbol{J}$ that has the same global statistics as $\boldsymbol{I}$ but with a greater correlation with $\boldsymbol{P}$ will produce superior evaluation metric scores in component substitution pan-sharpening methods. For our image pre-processing, pan-sharpening and evaluation we use the toolbox ‘Pansharpening Tool ver 1.3’, created by Vivone et al. (2015). This toolbox can be downloaded for free at `http://openremotesensing.net/knowledgebase/a-critical-comparison-among-pansharpening-algorithms`.

7.2.1 Image Acquisition and Pre-Processing

We acquired 16-bit multispectral ($\boldsymbol{M}$) and panchromatic ($\boldsymbol{P}$) imagery from the Digital-Globe (Maxar) sample series found at `https://www.digitalglobe.com/samples`. This imagery consists of four datasets: Washington DC (2048 x 2048), Stockholm (512 x 512), Tripoli (512 x 512) and Rio de Janeiro (512 x 512). The spatial resolution of $\boldsymbol{P}$ is four times greater than that of $\boldsymbol{M}$ for all of the datasets. The toolbox contains functions for image viewing (‘viewimage’) and downsampling (‘resize_image’). Using the viewimage function, we exported the normalised datasets ($\boldsymbol{M}$ and $\boldsymbol{P}$) to use as the imagery in our experiment. Storing a copy of the original $\boldsymbol{M}$ to be used as a ground-truth ($\boldsymbol{T}$), we followed the scheme outlined in Figure. 7.1 to produce lower resolution imagery. In Figure. 7.1, ‘Reduce’ refers to downsampling that was carried out using the ‘resize_image’ function. This function and takes both $\boldsymbol{T}$ and $\boldsymbol{P}^*$ as inputs and produces $\boldsymbol{P}$ and $\boldsymbol{M}^*$; degraded version that have been reduced in resolution by a pre-defined integer (4 in our case). To bring $\boldsymbol{M}^*$ back to the same size as $\boldsymbol{T}$ we perform upsampling (interpolate) using the standard bi-cubic interpolator to produce $\boldsymbol{M}$. With spatially degraded imagery to pan-sharpen and a ground-truth to act as a basis for comparison, we are now equipped to apply Wald’s protocol post pan-sharpening.

Out of the four Digital-Globe (Maxar) datasets, Washington DC and Stockholm were captured using the World-View 2 (WV-2) satellite whereas Tripoli and Rio de Janeiro used the more recent World-View 3 (WV-3). Both of these satellites collect radiation from eight multispectral bands with the WV-3 having eight additional short-wave-infrared sensors. As our method for linear mapping only allows for four bands we limit our choice to blue, green, red and near-infrared

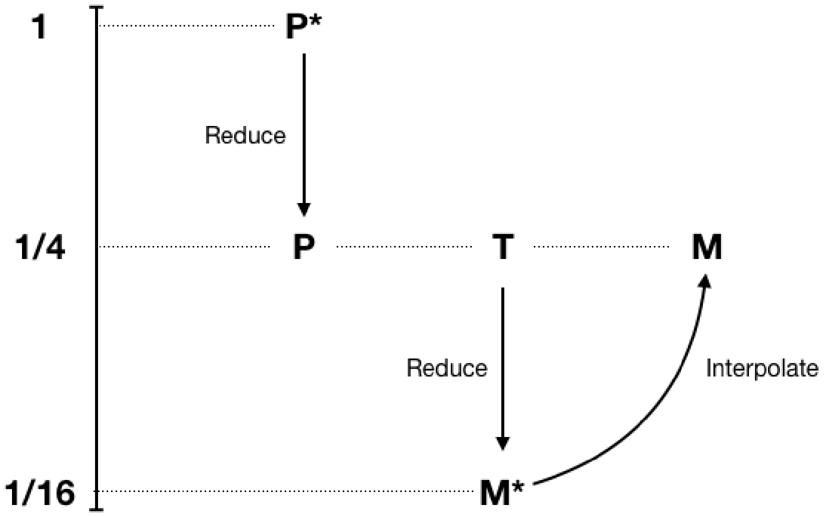


Figure 7.1: Image pre-processing for Wald's protocol. We produce the ground-truth ($\boldsymbol{T}$) and degraded panchromatic ($\boldsymbol{P}$) and spectral bands ($\boldsymbol{M}$). The fractions on the left refer to the ratio in size with the original panchromatic image ($\boldsymbol{P^*}$).

1, a decision that was based on what earlier satellites were restricted to (IKONOS, QuickBird).

The toolbox comes with the dataset Toulouse (1024 x 1024) that has been captured by the French Space Agencies Pléiades aerial platform. This dataset comes without a panchromatic image so the authors instead synthesised their own using the red and green bands, a series of filtering steps and a nominal MTF for the panchromatic sensor.

The toolbox contains the gain values at Nyquist used in the MTF for several satellites, one of which is the WV-2. Upon contacting the authors we were supplied the gain values needed for the WV-3 imagery which we incorporated into their toolbox. These values are required in the toolboxes 'resize_images' function which generates our lower resolution imagery needed for Wald's protocol to be tested.

As we described in Chapter 6, CS pan-sharpening methods usually require statistically matching the panchromatic image to the component it is replacing i.e. the intensity mapping.

To match the first two statistical moments of $\boldsymbol{P}$ to $\boldsymbol{I}$ it is common (Dou and Chen, 2008, Rahmani et al., 2010, Vivone et al., 2015) to use the linear relationship

$$\boldsymbol{P}_{sm} = \frac{\sigma_{\boldsymbol{I}}}{\sigma_{\boldsymbol{P}}}(\boldsymbol{P} - \mu_{\boldsymbol{P}}) + \mu_{\boldsymbol{I}}, \tag{7.5}$$

where $\boldsymbol{P}_{sm}$, $\boldsymbol{P}$, $\mu_{\boldsymbol{P}}$, $\mu_{\boldsymbol{I}}$, $\sigma_{\boldsymbol{P}}$ and $\sigma_{\boldsymbol{I}}$ are the statistically matched panchromatic and original panchromatic image respectively. Mean centering $\boldsymbol{P}$ by subtracting its mean ($\mu_{\boldsymbol{P}}$) allows for a controlled way of altering the panchromatic images standard deviation ($\sigma_{\boldsymbol{P}}$). If the multiplier for ($\boldsymbol{P} - \mu_{\boldsymbol{P}}$) is the ratio of the intensity images standard deviation ($\sigma_{\boldsymbol{I}}$) to the $\sigma_{\boldsymbol{P}}$ then we have matched the variance of $\boldsymbol{P}$ to $\boldsymbol{I}$. The addition of a constant to this controls the new mean of our output image; choosing the mean of $\boldsymbol{I}$ ($\mu_{\boldsymbol{I}}$) gives us our new panchromatic image which possesses the same means and and variance as that of $\boldsymbol{I}$.

Previously we have mentioned how statistically matching $\boldsymbol{I}$ to $\boldsymbol{P}$ reduces the chromatic differences between $\boldsymbol{M}$ and our pan-sharpened image $\boldsymbol{H}$. An additional effect is that statistical matching $\boldsymbol{P}$ to $\boldsymbol{I}$ means that the edge map produced from their difference has a zero mean and will therefore not change the global brightness of $\boldsymbol{H}$ from that of $\boldsymbol{M}$ post pan-sharpening.

7.2.2 Methods

With respect to $\boldsymbol{I}$, our aim is to report on whether using our greyscale mapping $\boldsymbol{J}$ which has a greater correlation with the panchromatic image $\boldsymbol{P}$, than the standard intensity mapping $\boldsymbol{I}$, produces superior evaluation scores. With this in mind, we have selected the methods that allow for CS of our mapping $\boldsymbol{J}$ with the intensity mapping $\boldsymbol{I}$:

- Generalised Intensity-Hue-Saturation (GIHS) Tu et al. (2001)
- Brovey Transform (BT) Gillespie et al. (1987)
- Gram-Schmidt mode 1 (GS1) Laben and Brower (2000)

For the sake of comparison we have also included results from the methods:

- Gram-Schmidt mode 2 (GS2) Laben and Brower (2000)
- Principle Component Analysis (PCA) Chavez et al. (1991)
- Band Dependent Spatial Detail (BDSD) Garzelli et al. (2008)
- Partial Replacement Adaptive Component Substitution (PRACS) Choi et al. (2011)
- High Pass Filtering (HPF) Chavez et al. (1991)

- Smoothing Filter Based Intensity Modulation (SFIM) Liu (2000)

These methods have been discussed in Chapter 5. The first five methods comprise the classical CS techniques. BDSD and PRACS are the latest CS methods available in the toolbox and will provide a reference as to how well the classical techniques that use $\boldsymbol{J}$ score. The last two methods in the second list contain filter based approaches which we decided to include because they are more selective at extracting high-frequency information from $\boldsymbol{P}$ and don't require complex transforms i.e. wavelets.

In the next section we shall present the imagery that have been pan-sharpened using these methods.

7.3 Results and Discussion

Figures. 7.2 and 7.4 show the World-View 3 datasets; Rio de Janeiro and Tripoli, Figures. 7.6 and 7.8 show the World-View 2 datasets; Stockholm and Washington DC. Lastly, Figure. 7.10 shows the Pléiades dataset; Toulouse. To assist in demonstrating the pan-sharpening effects of each algorithm we have included Figures. 7.3, 7.5, 7.7, 7.9 and 7.11. These figures show where we have manually zoomed-in on the initial pan-sharpened images to better show the effects of the sharpening.

Each of these figures show the ground-truth ($\boldsymbol{T}$), its degradation ($\boldsymbol{M}$) and their outputs from the various pan-sharpening algorithms. Where the (+) notation has been used for the methods IHS, BT and GS defines where we have used our greyscale mapping $\boldsymbol{J}$ instead of $\boldsymbol{I}$. To show the differences between $\boldsymbol{J}$ and $\boldsymbol{I}$ we have, in a similar manner to Chapter 6, included the edge map produced from $\boldsymbol{P}-\boldsymbol{J}$ and its difference with the edge map $\boldsymbol{P}-\boldsymbol{I}$

$$\frac{\boldsymbol{P}-\boldsymbol{J}}{2\boldsymbol{P}-\boldsymbol{I}-\boldsymbol{J}}. \tag{7.6}$$

Computing the ratio of the edge maps in this way means that when $\boldsymbol{I}=\boldsymbol{J}$ then the calculation reduces to produce a greyscale pixel value of 0.5. Instances when $\boldsymbol{P}=\boldsymbol{J}$ will result in a black pixel and conversely a white pixel is produced when $\boldsymbol{P}=\boldsymbol{I}$.

In Chapter 6, we saw for the dataset Tripoli the closeness in the coefficients that map $\boldsymbol{M}$

to $\boldsymbol{J}$ with respect to $\boldsymbol{I}$ and their near identical correlation with $\boldsymbol{P}$. We also saw how this led to identical edge maps between the two mappings when Eq. (7.6) was visualised (shown again in Figure. 7.4 where the image was dominated by pixel values at 0.5). In Figure. 7.5 we show where we have zoomed-in on the image and this allows us to see slight noise in the image. For the remaining difference map images we see stronger differences in the form of image feature edge definition.

The bottom three rows of Figures. 7.2 to 7.10 show the outputs ($\boldsymbol{H}$) from the various pan-sharpened methods. Upon initial visual inspection all of the methods provide improved spatial resolution compared to that of the degraded spectral image $\boldsymbol{M}$. It is not until we look at enlarged areas of these figures, (Figures. 7.3, 7.5, 7.7, 7.9 and 7.11) do we begin to see disparity in the visual quality of the methods. The CS methods (that use either $\boldsymbol{I}$ or $\boldsymbol{J}$) still boast enhanced visual spatial resolution, however methods that incorporate the bi-cubic interpolator in their fusion process i.e. PRACS suffer dramatically with blurring. Khan et al. (2008) have reported that bi-cubic interpolation is prone to this introduction of image blur. Another cause of image blur has by attributed to poor registration between the $\boldsymbol{P}$ and $\boldsymbol{M}$ which, in the case of SFIM, produces blurring when the high-frequency information of $\boldsymbol{P}$ is extracted when they ratio it with the low-pass filtered $\boldsymbol{P}$.

It has been reported (González-Audícana et al., 2004) that relative to the multi-resolution bracket of pan-sharpening methods, CS techniques fail to preserve the spectral integrity of the original unsharpened images. We can visually see examples of this by again looking at Figure. 7.3 where the colour of the red object in the north-east quadrant has been lost in the IHS and BT methods and is only vaguely evident with PCA and GS. Visually, SFIM comes close to preserving the spectral content of this object. The problem with this method is that it relies on computing a ratio (similarly with the BT), and this can sometimes cause excessive pixel modulation values resulting in unnatural levels of brightness, see the sports field in Figure. 7.9. This image of Washington DC also causes problems for the PCA method where there are very noticeable areas of contrast inversion (1) the sports field appears almost black relative to the forestry and (2) the area of grass at the bottom of the image is overly bright.

Visually the BDSD method excels at adding high levels of spatial detail while maintaining the spectral integrity of the original $\boldsymbol{M}$ bands. However it is not until we look into the evaluation

Table 7.1: World-View 3.

(a) Rio de Janeiro

Method	Q4	SAM	ERGAS
GIHS	0.7271	3.8076	7.0932
IHS+	0.7277	3.7527	7.0939
BT	0.7214	3.7324	7.1336
BT+	0.7229	3.7307	7.0916
GS	0.7273	3.8966	7.1149
GS+	0.7280	3.8101	7.1153
PCA	0.7265	3.9773	7.1209
PRACS	0.7073	3.4009	7.4658
GS2-GLP	0.7253	3.4604	7.7132
BDSD	0.7110	3.6822	7.9963
HPF	0.7230	3.4768	7.5790
SFIM	0.7244	3.3573	7.5218

(b) Tripoli

Method	Q4	SAM	ERGAS
GIHS	0.8817	5.9491	6.1575
IHS+	0.8817	5.9495	6.1574
BT	0.8755	5.9231	6.2405
BT+	0.8749	5.9142	6.2363
GS	0.8817	5.8615	6.1319
GS+	0.8818	5.8618	6.1318
PCA	0.8820	5.8619	6.1229
PRACS	0.8652	5.1664	6.3958
GS2-GLP	0.8867	5.3487	6.4732
BDSD	0.8939	5.3632	6.4546
HPF	0.8692	5.3902	6.5385
SFIM	0.8695	5.4381	6.4896

scores that are based in making comparisons with the ground-truth do we get an objective measure of how well it and every other method does. Tables. 7.1 to 7.3 show the evaluation scores for the images acquired by the three satellites: World-View 3, World-View 2 and Pléiades respectively. The ideal evaluation metric scores that each method aspires to achieve for Q4, SAM and ERGAS is 1, 0 and 0 respectively. The evaluation scores for the image Rio de Janeiro are shown in Table. 7.1(a) where we can see that the GS+ method ranks the highest for the Q4 metric. This metric is grounded in the global statistics of the image, which given the large correlation that $\boldsymbol{I}$ and $\boldsymbol{J}$ have with $\boldsymbol{P}$, has meant that it and other CS methods have scored highly. The BT+ has the best ERGAS score, which we again attribute to the high correlation of $\boldsymbol{J}$ with $\boldsymbol{P}$ and with its band dependent injection gains. Band dependent gains play an important role to avoid saturation distorting of image pixels and these methods tend to have a superior SAM score.

For the Tripoli dataset (Table. 7.1(b)), the BDSD scores the highest for the Q4 metric followed by the CS methods. Of these our $\boldsymbol{J}$ mapping does not produce a significant score difference to the classic techniques; a consequence of $\boldsymbol{J} \approx \boldsymbol{I}$ for this image. PRACS obtains the best SAM score and the classic techniques beat the remaining methods on ERGAS.

As we mentioned earlier, the PCA and SFIM methods cause severe distortions for the Washington DC dataset (Table. 7.2(b)). This reflects in their metric scores where SFIM has scores

Table 7.2: World-View 2.

(a) Stockholm

Image	Q4	SAM	ERGAS
GIHS	0.7458	7.8753	8.4829
IHS+	0.7460	7.8891	8.4795
BT	0.7511	7.1486	8.4366
BT+	0.7509	7.1467	8.3959
GS	0.7470	7.8452	8.4309
GS+	0.7471	7.8587	8.4265
PCA	0.7371	8.3112	8.5701
PRACS	0.7228	6.8199	9.0781
GS2-GLP	0.7346	6.7708	9.2901
BDSD	0.7244	7.0330	9.8059
HPF	0.7365	6.7901	9.1560
SFIM	0.7520	6.0228	8.8909

(b) Washington DC

Method	Q4	SAM	ERGAS
GIHS	0.7697	12.1000	9.6328
IHS+	0.7697	12.1110	9.6240
BT	0.7580	11.3424	10.0189
BT+	0.7586	11.3194	9.9816
GS	0.7166	13.0366	10.2719
GS+	0.7162	13.0637	10.2663
PCA	0.4718	21.3630	18.1914
PRACS	0.6717	11.4194	11.6626
GS2-GLP	0.6565	16.1856	14.4686
BDSD	0.6996	17.7455	15.5249
HPF	0.7165	15.1197	12.5761
SFIM	0.3801	17.0674	448330

that are so poor (ERGAS 448330) as to make it incomparable to the other methods. As with the BT method, the SFIM extracts the high-frequency components by taking a ratio. In particular, the SFIM method relies on dividing $\boldsymbol{P}$ with a blurred copy. During this resolution reduction it is possible that pixel values are assigned to approximate zero, and this means that it can cause excessively large values in the modulation image. Effects of this can be seen in Figure. 7.7 (SFIM), where we can see overly bright pixel values in and around the baseball field, and are the cause the excessively large metric result in Table. 7.2. The highest scoring methods for this dataset are the IHS and BT followed by PRACS.

Although the CS methods score the lowest for the Q4 metric on the Pléiades dataset (Table. 7.3), our $\boldsymbol{J}$ mapping increases the performance of the methods against the classics by the largest margin when compared with the other datasets. The SAM and ERGAS metrics see the BT+ and GS+ having the best scores respectively.

As we have already mentioned, the Pléiades dataset comes with the toolbox 'Pansharpening Tool ver 1.3' and is presented in the publication by Vivone et al. (2015). Upon comparison of our metric scores with theirs, we see that our scores are significantly worse. We attribute this to a variation in the experimental methods. The method that Vivone et al. (2015) follows is based on pan-sharpening then evaluating non-normalised data prior to visualising it with the 'viewimage' function (which contrast stretches to fill the range of the display subject to clipping in the upper

Table 7.3: Pléiades.

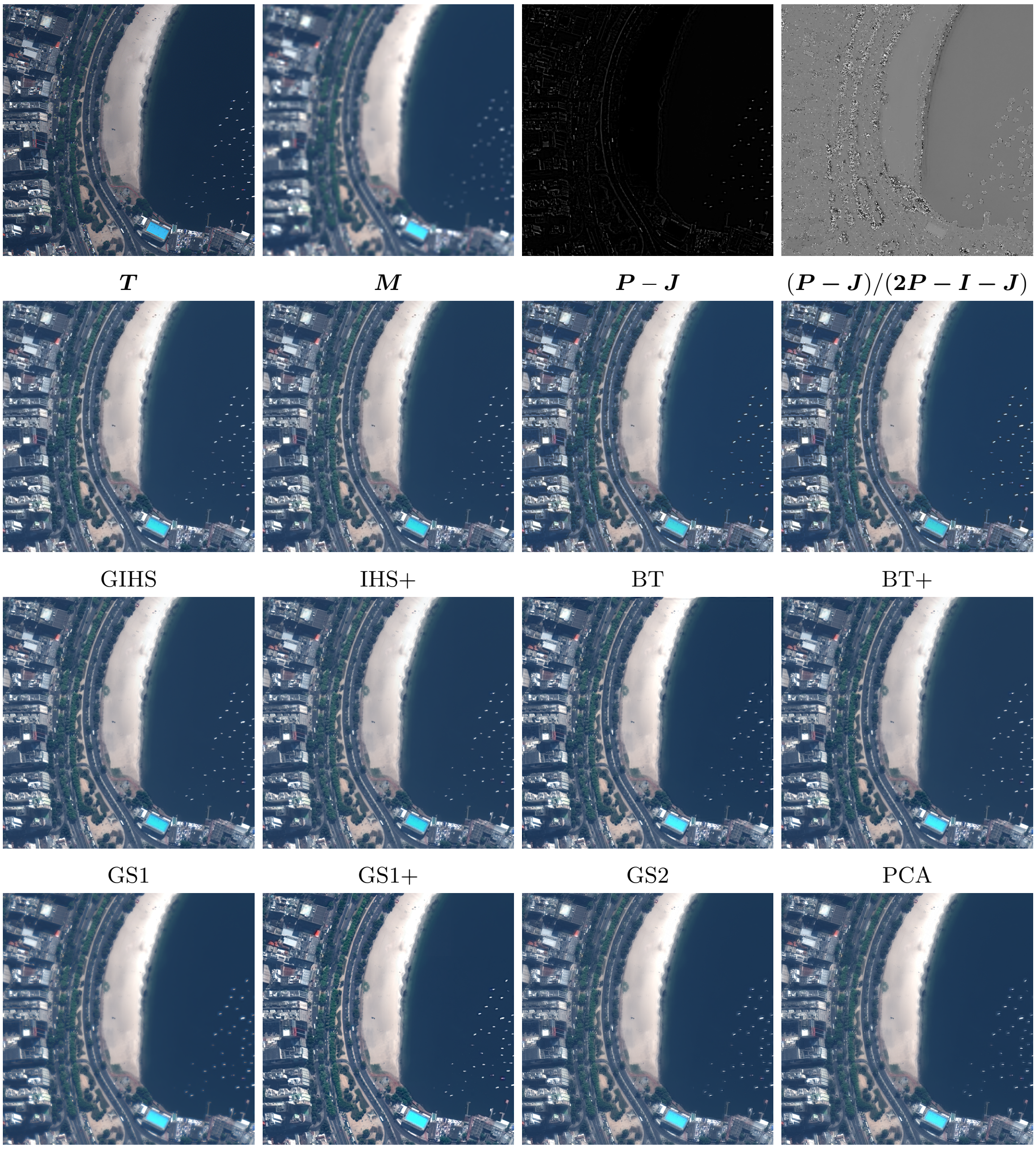
(a) Toulouse

Method	Q4	SAM	ERGAS
GIHS	0.8922	8.0786	7.2902
IHS+	0.8936	8.0564	7.2495
BT	0.8915	7.4108	7.4934
BT+	0.8918	7.3984	7.4561
GS	0.8918	8.0105	7.2871
GS+	0.8933	7.9754	7.2423
PCA	0.9068	7.7823	6.9154
PRACS	0.9184	7.6598	7.3829
GS2-GLP	0.9096	10.007	8.3339
BDSD	0.9052	9.8802	8.5868
HPF	0.9222	7.9789	7.6064
SFIM	0.9227	7.6274	7.5696

and lower quantile). This is in opposition to our method where we contrast stretched our images as a pre-processing step. This introduces pixel clipping for the upper and lower quantiles of pixel values in the image with the result being that they be clipped to either 1 or 0 during the export procedure and thusly effecting the metric scores.

Lastly we have included Table. 7.4 which contains the relative metric placement ranks for each method over all datasets. If a particular method has the best metric score (accurate to 4 decimal places) relative to the rest, then it was ranked first. For example, for the Washington DC dataset, the methods GIHS and IHS+ both had the highest Q4 values and as a result were given a first place rank for that metric. SFIM had the lowest Q4 score and was therefore ranked last for the dataset. Ideally a method aspires to achieve a first place rank for all metrics on all datasets. The summation of these ranks in this case would give an ideal total of 15. The last column (Rank) in Table. 7.4 shows these rank summations and their placement rank (value inside the parenthesis) relative to each other. Displaying the results in this way allows us to show the overall performance of the methods in our experiment in a manner where outliers (see the Washington DC Q4 value for the SFIM method in Table. 7.2) from their metric scores won't severely effect their overall rank placement.

From Table. 7.2 we can see that across all of the datasets, our $\boldsymbol{J}$ mapping improves the classical pan-sharpening methods to an extent that they take the top three positions.



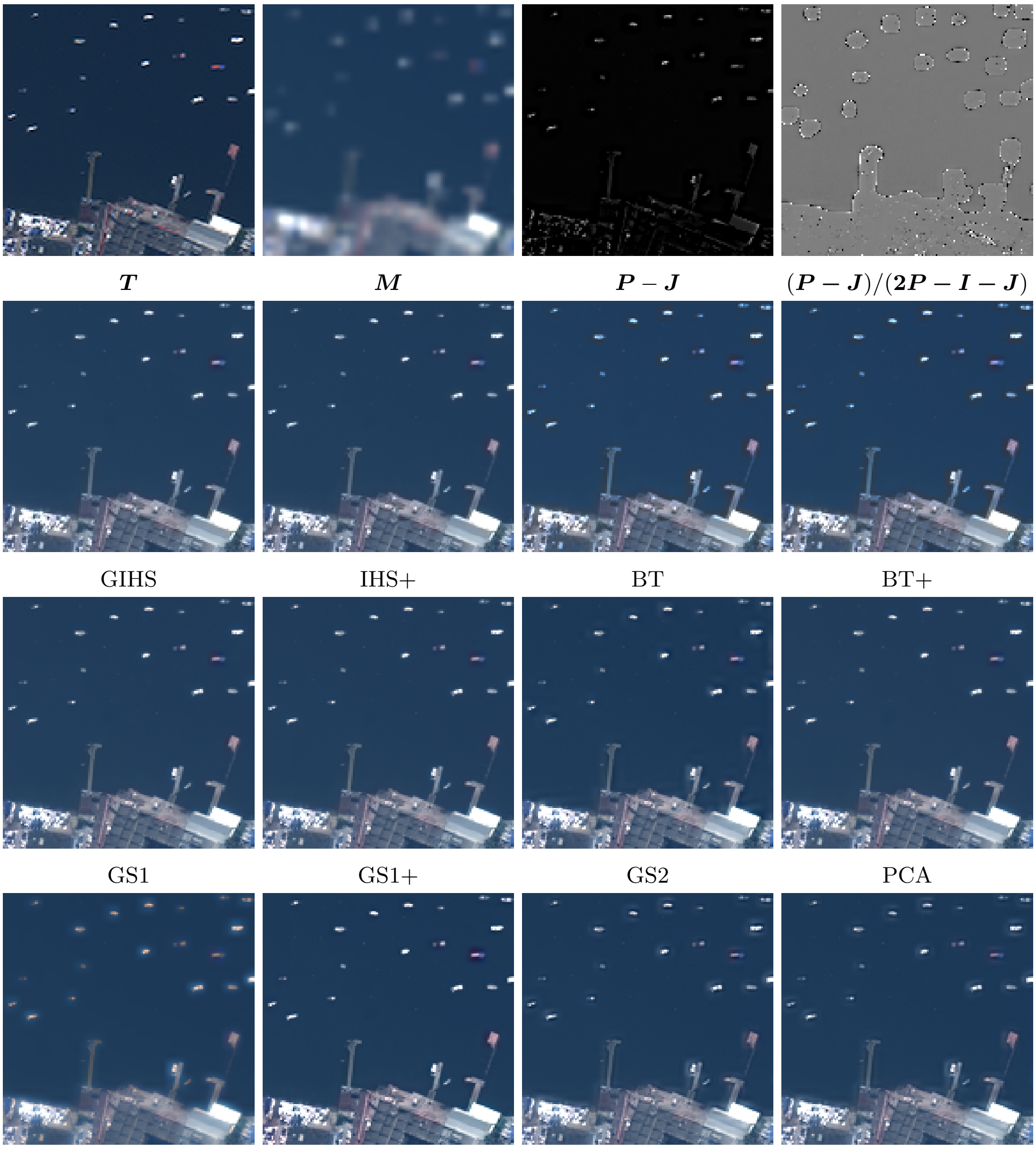
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$

GIHS IHS+ BT BT+

GS1 GS1+ GS2 PCA

PRACS BDSD HPF SFIM

Figure 7.2: World-View 3, Rio de Janeiro. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



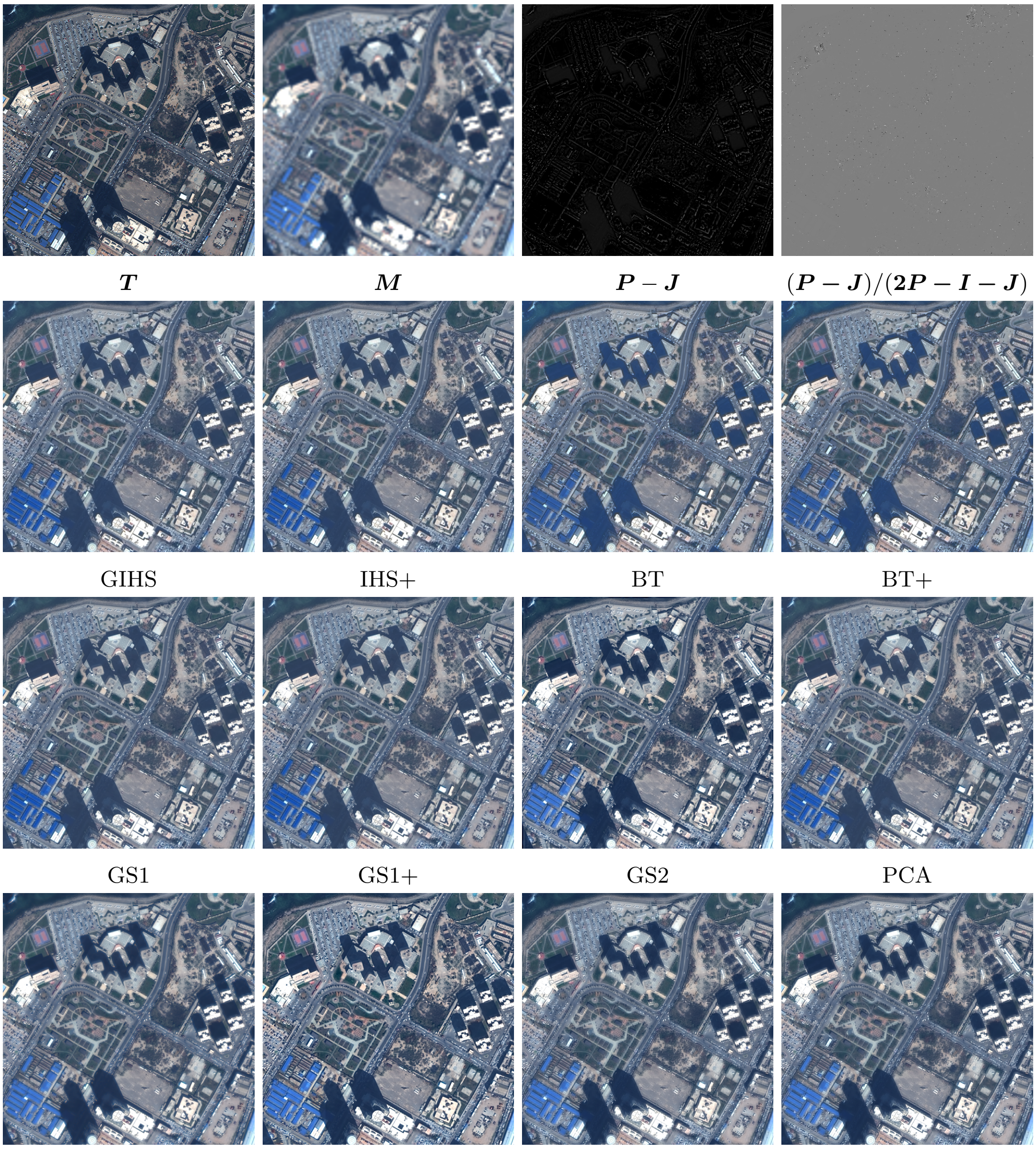
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$

GIHS IHS+ BT BT+

GS1 GS1+ GS2 PCA

PRACS BDSD HPF SFIM

Figure 7.3: World-View 3, Rio de Janeiro zoomed-in. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



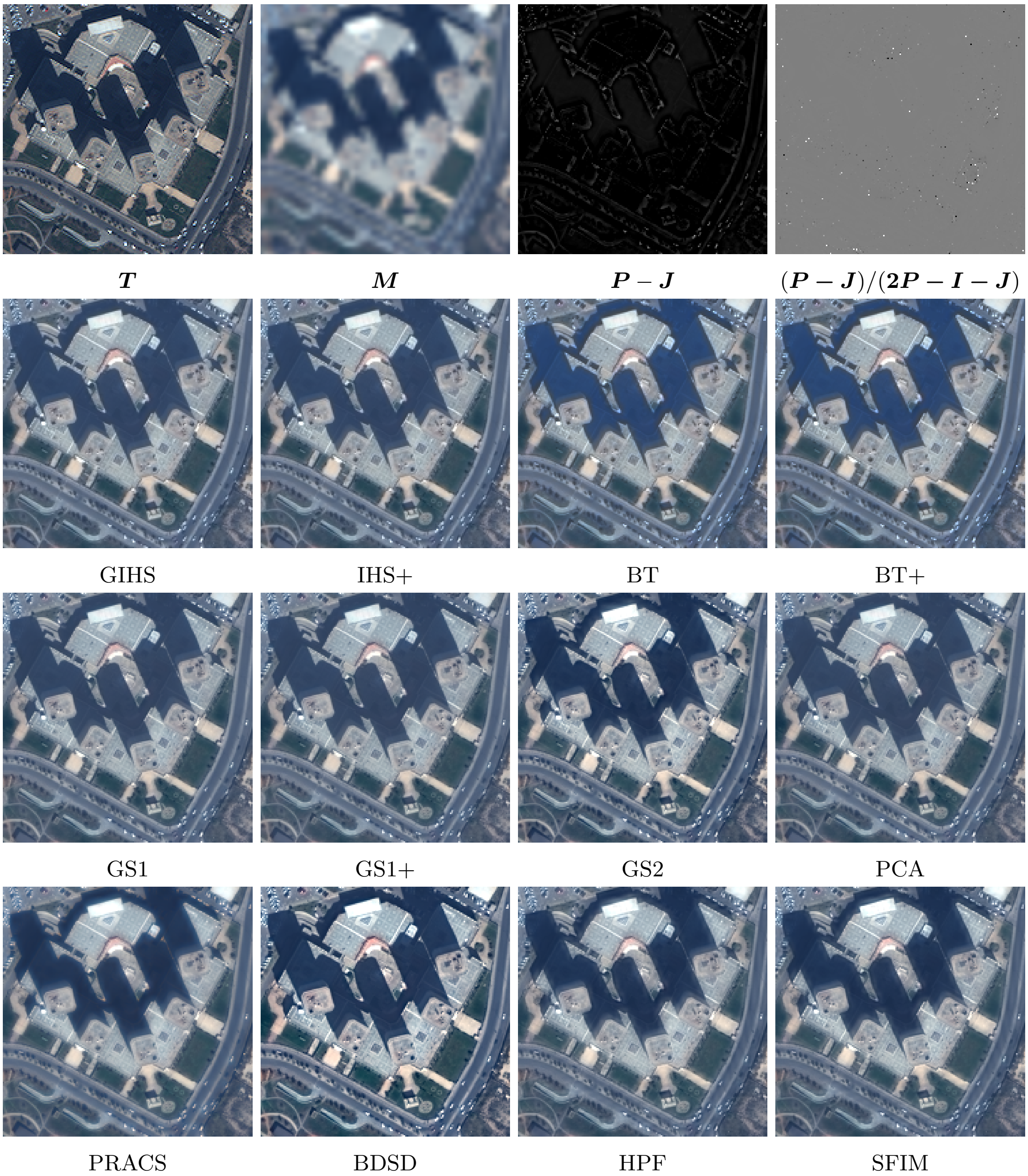
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$

GIHS IHS+ BT BT+

GS1 GS1+ GS2 PCA

PRACS BDSD HPF SFIM

Figure 7.4: World-View 3, Tripoli. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



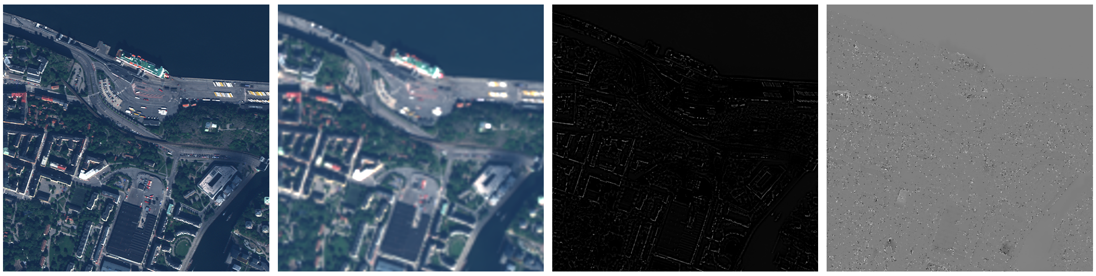
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$

GIHS IHS+ BT BT+

GS1 GS1+ GS2 PCA

PRACS BDSD HPF SFIM

Figure 7.5: World-View 3, Tripoli, zoomed-in. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$

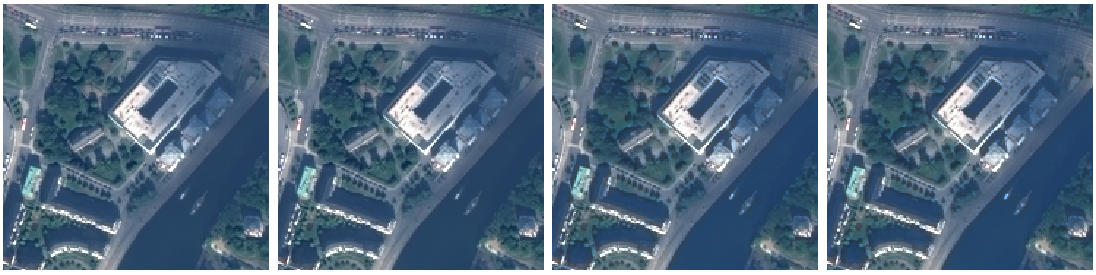
GIHS IHS+ BT BT+

GS1 GS1+ GS2 PCA

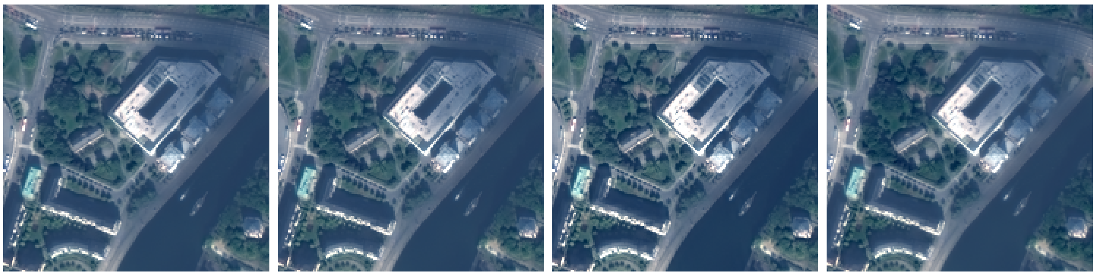
PRACS BDSD HPF SFIM

Figure 7.6: World-View 2, Stockholm. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.

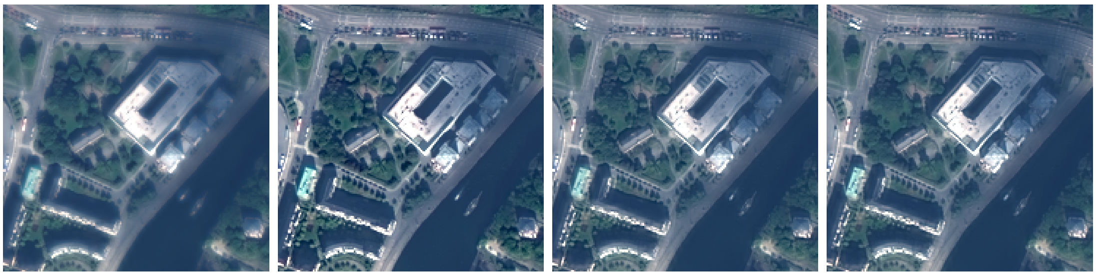
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$



GIHS IHS+ BT BT+

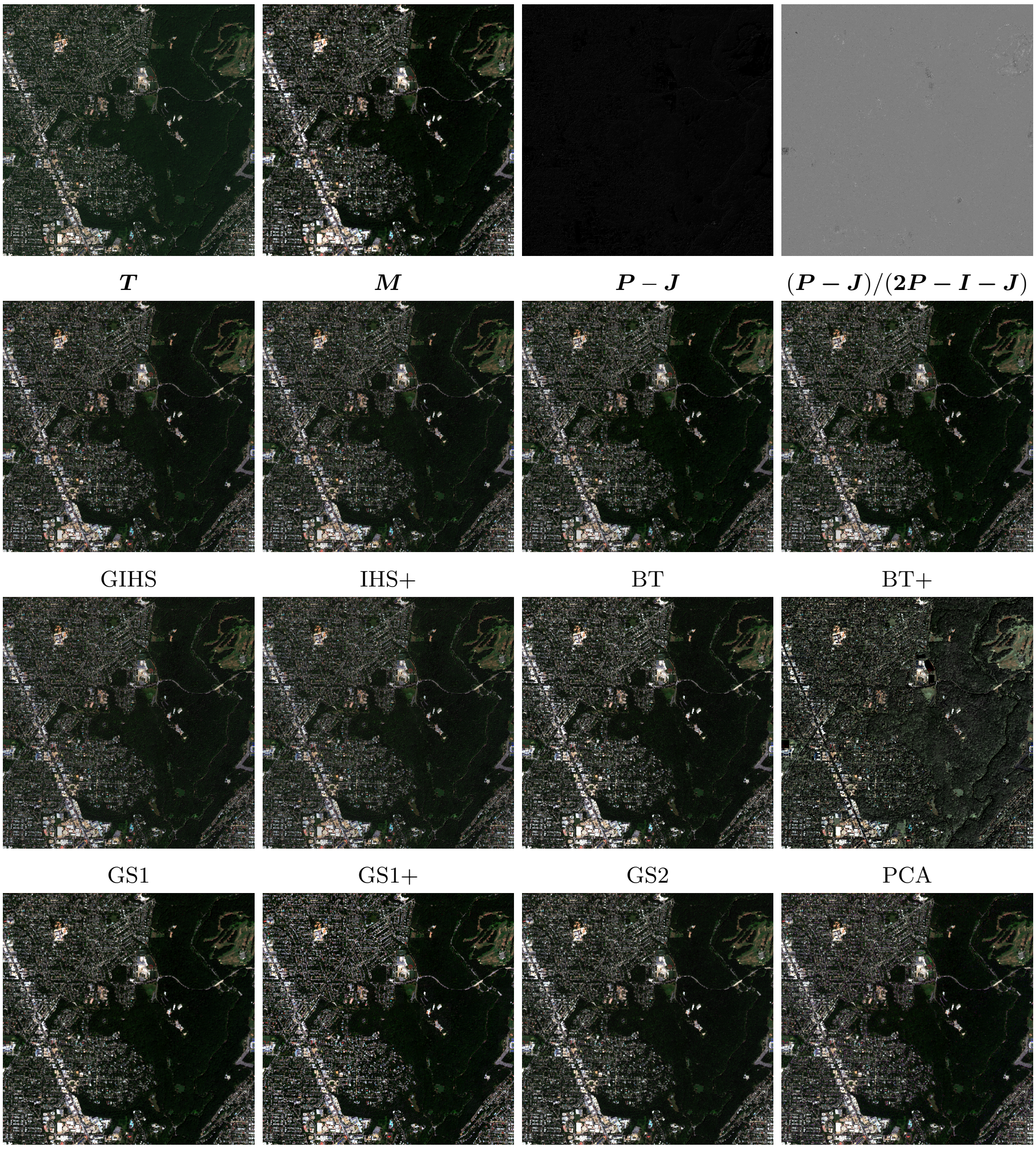


GS1 GS1+ GS2 PCA



PRACS BDSD HPF SFIM

Figure 7.7: World-View 2, Stockholm, zoomed-in. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



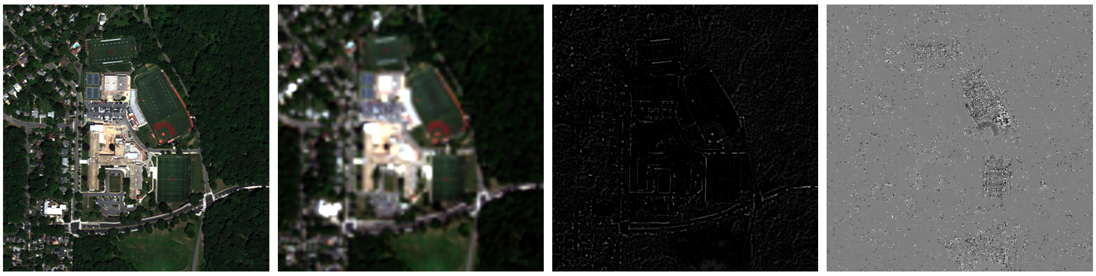
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $(\boldsymbol{P}-\boldsymbol{J})/(\boldsymbol{2P}-\boldsymbol{I}-\boldsymbol{J})$

GIHS IHS+ BT BT+

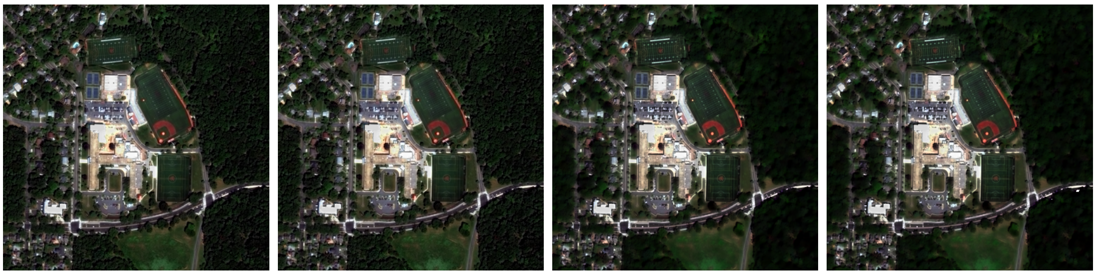
GS1 GS1+ GS2 PCA

PRACS BDSD HPF SFIM

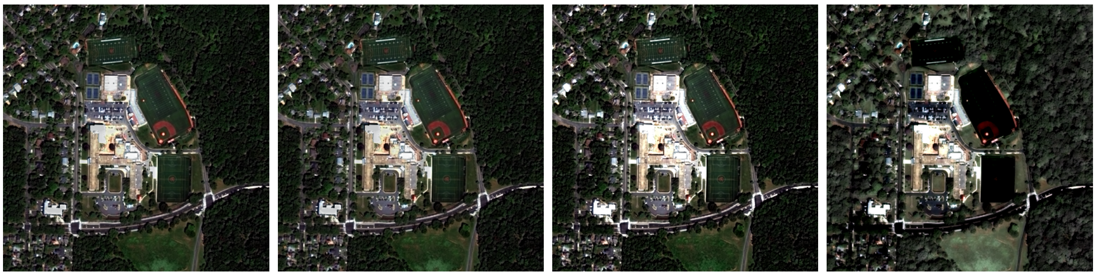
Figure 7.8: World-View 2, Washington DC. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



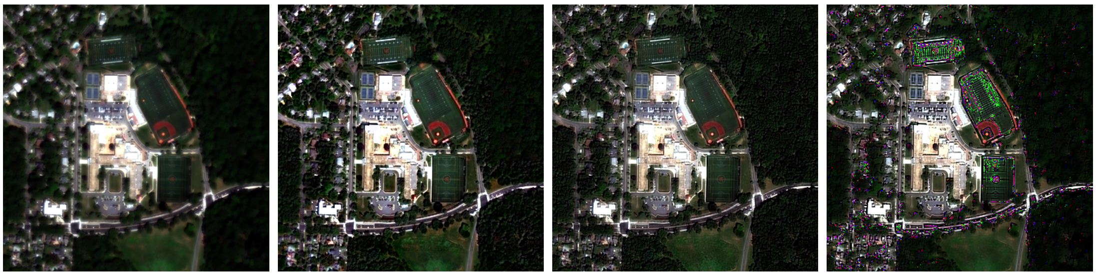
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P-J}$ $\boldsymbol{P-J/2P-I-J}$



GIHS IHS+ BT BT+

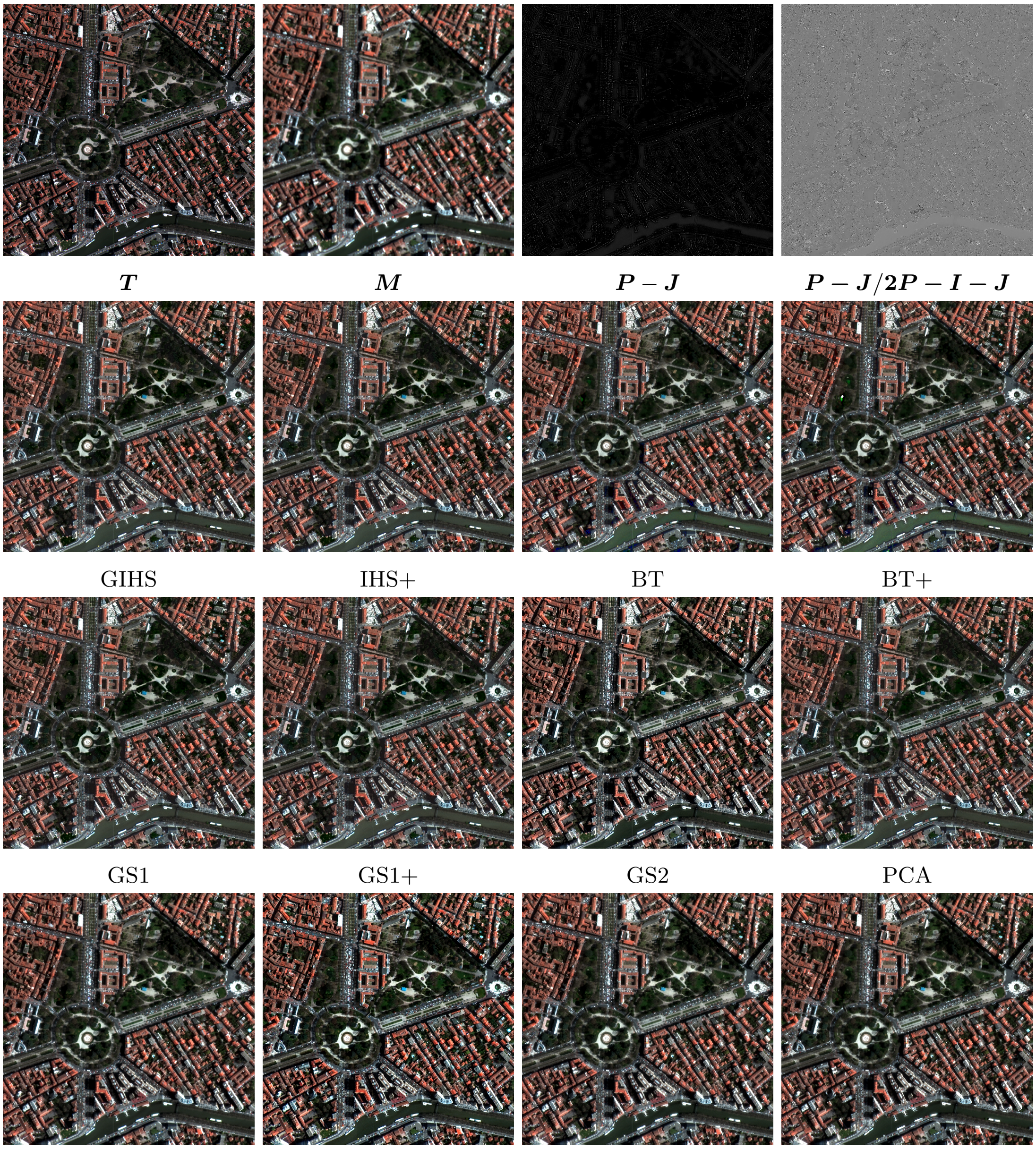


GS1 GS1+ GS2 PCA



PRACS BDSD HPF SFIM

Figure 7.9: World-View 2, Washington DC, zoomed-in. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P-J}$ and its normalised ratio with $\boldsymbol{P-I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



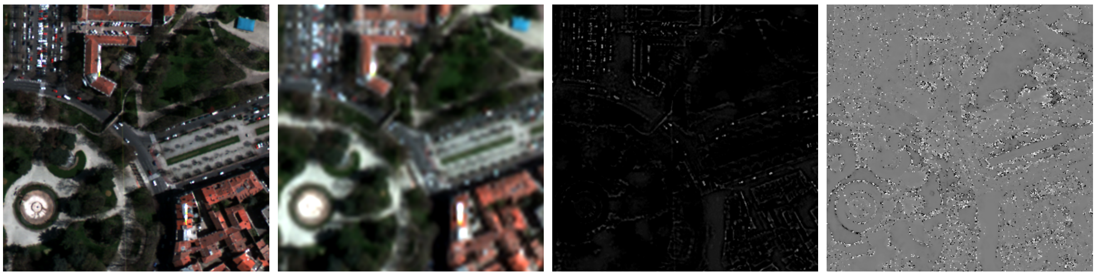
$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P}-\boldsymbol{J}$ $\boldsymbol{P}-\boldsymbol{J}/2\boldsymbol{P}-\boldsymbol{I}-\boldsymbol{J}$

GIHS IHS+ BT BT+

GS1 GS1+ GS2 PCA

PRACS BDSD HPF SFIM

Figure 7.10: Pléiades, Toulouse. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P}-\boldsymbol{J}$ and its normalised ratio with $\boldsymbol{P}-\boldsymbol{I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.



$\boldsymbol{T}$ $\boldsymbol{M}$ $\boldsymbol{P-J}$ $\boldsymbol{P-J/2P-I-J}$

GIHS IHS+ BT BT+

GS1 GS1+ GS2 PCA

PRACS BDSD HPF SFIM

Figure 7.11: Pléiades, Toulouse zoomed-in. First row: ground-truth ($\boldsymbol{T}$), multispectral image $\boldsymbol{M}$, edge map $\boldsymbol{P-J}$ and its normalised ratio with $\boldsymbol{P-I}$. Last three rows: pan-sharpened images $\boldsymbol{H}$.

Table 7.4: Metric ranks for each method across the five datasets.

Method	Metric	Rio de Janeiro	Tripoli	Stockholm	Wash. DC	Toulouse	Rank
GIHS	Q4	4	5	7	1	9	90 (5)
	SAM	9	11	10	4	10	
	ERGAS	2	5	6	2	5	
IHS+	Q4	2	5	6	1	7	82 (3)
	SAM	8	12	11	5	9	
	ERGAS	3	4	5	1	3	
BT	Q4	10	6	2	3	11	90 (5)
	SAM	7	10	7	2	2	
	ERGAS	7	7	4	4	8	
BT+	Q4	9	7	3	2	10	72 (1)
	SAM	6	9	6	1	1	
	ERGAS	1	6	1	3	7	
GS	Q4	3	5	5	4	10	86 (4)
	SAM	11	6	8	6	8	
	ERGAS	4	3	3	6	4	
GS+	Q4	1	4	4	6	8	78 (2)
	SAM	10	7	9	7	6	
	ERGAS	5	2	2	5	2	
PCA	Q4	5	3	8	10	5	106 (7)
	SAM	12	8	12	12	5	
	ERGAS	6	1	7	11	1	
PRACS	Q4	12	10	12	8	3	97 (6)
	SAM	2	1	4	3	4	
	ERGAS	8	8	9	7	6	
GS2-GLP	Q4	6	2	10	9	4	111 (9)
	SAM	3	2	2	9	12	
	ERGAS	11	10	11	9	11	
BDSD	Q4	11	1	11	7	6	126 (10)
	SAM	5	3	5	11	11	
	ERGAS	12	9	12	10	12	
HPF	Q4	8	9	9	5	2	109 (8)
	SAM	4	4	3	8	7	
	ERGAS	10	12	10	8	10	
SFIM	Q4	7	8	1	11	1	97 (6)
	SAM	1	5	1	10	3	
	ERGAS	9	11	8	12	9	

7.4 Conclusion

In this chapter we reported on the effect of using a greyscale ($\boldsymbol{J}$), that preserved the global statistics of the intensity mapping ($\boldsymbol{I}$), while having a greater correlation with the panchromatic ($\boldsymbol{P}$) had on component substitution pan-sharpening methods.

Over the course of the experiment we found that over a total of five datasets and three evaluation metrics, substituting $\boldsymbol{J}$ into the intensity-hue-saturation (IHS+), Brovey transform (BT+) and Gram Schmidt (GS+) led to superior evaluation scores. Upon ranking the summation of the ranked metric scores across the datasets showed that these enhanced classical techniques outperformed the more recent competing methods in our experiment, with the BT+ being ranked first followed by the GS+ and IHS+. From this we conclude that our solution can be used in the classical CS methods to outperform more recent CS methods despite the numerous constraints that prioritise preserving the global statistics over correlation on our computation of $\boldsymbol{J}$.

The more recent CS methods that used the bi-cubic interpolator were seen to score poorly which we conclude is due to the introduction of misinformation into the spectral bands during these sampling steps.

Further work on this study would look at the impact of removing the constraint on variance to see how much it had inhibited the potential increase in correlation, and how this would effect the evaluation metric scores. Additionally we would be interested in looking at the effects of changing certain image-preprocessing steps; like looking at linearly scaling the image post-sharpening instead of prior to sharpening (which introduced pixels being clipped to 0 or 1).

Chapter 8

Conclusions and Further Work

This chapter consists of two sections. The first section concludes the total body of work presented in this thesis by highlighting our contributions made throughout. The second section discusses further work that could improve upon our current solution and bring to light alternative applications to problems in imaging.

8.1 Conclusion

In the field of imaging, linear mappings have the simplest mathematical model and offer the lowest time complexity in producing greyscales from multiband imagery. It has found use in 1) colour-to-greyscale, where the aim is to reproduce an image that best represents colour contrast and 2) panchromatic image sharpening, where the desire is to create a greyscale that has maximum positive correlation with the higher resolution panchromatic image. The desire of colour-to-grey is still popular for vintage style imagery and a cheaper alternative to colour printing whereas in multiband-to-grey, the image is used as a low-pass approximation to create an unsharp mask with the panchromatic image.

Depending on the desired properties of the greyscale is the deciding factor on how weights are chosen for linearly mapping a multiband image into a single band. In colour-to-grey one of the challenges is avoiding the occurrence of mapping different colours with identical brightness values to the same shade of grey. In Chapter 3, we minimise the chances of this happening by optimising a constrained positive-definite quadratic objective function that produces band weights for a greyscale with maximum global variance. The convex set formed from these constraints led us to show that the optimum solution to our linearly constrained objective function for any positive-definite quadratic form will always exist at one of its vertices. The problem we then tackled was how we go about finding these vertices. While constraining our solution for the band weights to to lie within a unit cube as well as preserving the mean and energy of the colour image, we posed the solution using arithmetic operations on the colour images band means. This resulted in a sorting of three rational numbers to find the appropriate band weighting vector, an operation that can on average be solved with a time-complexity of $O(n)$.

An additional outcome from this optimisation occurs for colour images that do not possess full dynamic range and this will produce a greyscale image that similarly lacks full dynamic range. We addressed this scenario by releasing the cubic constraints on the band weights and instead imposed constraints on each pixel to ensure the final greyscale pixel values lay between 0 and 1. This problem became one of solving for an intersection of halfspaces in $\mathbb{R}$ and finding its vertices by number sorting: an operation that increased the time complexity of our solution to $O(n \log n)$. This solution to the colour-to-grey brightness preserving problem will always produce the optimal answer. Previous solutions to this optimisation used the Quadratic Programming function in MATLAB to solve which a) uses a higher time-complexity $O(n^3)$ and b) has been known to produce a greyscale that was not optimal.

All facets in the field of image fusion require evaluation metrics to assess the quality of the fused output image. In Chapter 4 we assessed the effectiveness of our brightness preserving colour-to-greyscale using psychophysics to analyse choices made in participant based preference experiments. Selecting three competing colour-to-grey techniques, our enhanced brightness preserving optimisation obtained a third place ranking across a standard image-set. The images where we came first typically possessed a low number of hues (~ 5) and it was reported that participants would penalise our greyscales for possessing unrealistic amounts of contrast.

Multispectral image fusion evaluation metrics adopt a more quantitative (and arguably simpler) approach to ranking fusion performance. Based on the rationale that if a greyscale produced from linearly mapping a multispectral image has a correlation of 1 with the panchromatic, then the statistical moments of both images would be the same. The addition of a quadratic constraint on image variance to our existing objective function (with the inclusion of an extra dimension to include the near-infrared) would therefore further limit the band weighting solution space for which to look for the solution to this ideal greyscale (if it exists). Chapter 6 addresses this problem in a unique and novel way by applying constrained multivariate calculus and linear algebra to the geometry produced by the objective function and the intersecting equality constraints. Fundamentally this involved us producing a novel solution to defining the equation of an ellipse that is produced by the intersection of two intersecting hyper-planes with a hyper-ellipse. By intersecting our convex set that contains the band weights for greyscales that fall within our desired dynamic range do we reduce our solution space to an elliptical line segment. Projecting the point of highest correlation (found by a least squares optimisation) to this segment do we find the weights for a greyscale that preserve our energy/mean and variance while increasing the correlation with the panchromatic image. Despite the increase (compared to our colour-to-greyscale) in arithmetic operations in producing our multispectral-to-greyscale, the time complexity is still limited by that of computing a convex hull.

We evaluated our pan-sharpened imagery using metrics based on the correlation coefficient and the root-mean-square error in association with Wald's Protocol. Using a published toolbox for pan-sharpening images, we incorporated our greyscale into the classical methods to test the hypothesis that preserving the statistical moments of the intensity mapping yet improving the correlation with the panchromatic will produce superior images when evaluated. Over five datasets and three metrics, our mapping produced a maximum increase of correlation with the panchromatic image of 0.15%. Ranking each dataset for each metric and then ranking their sum gave us an overall metric of performance for each method. From this we found that the three methods that used our greater correlated greyscale took the top three rankings: 1) Brovey Transform (BT+), 2) Gram-Schmidt (GS+) and 3) the Intensity-Hue-Saturation (IHS+).

8.2 Further Work

One of the advantages of the mathematical model presented in this thesis lies with its versatility. Our method of preserving the energy, mean, variance with further options on range and correlation make it applicable to many forms of dimensionality reduction problems. In the field of imaging, fusing near-infrared with the optical bands to produce high detail greyscales would find use in surveillance applications. In photography, fusing near-infrared with its respective colour image would improve clarity by compensating for the information lost by optical Rayleigh light scattering.

In direct relation to the work we presented in this thesis, we would begin with adapting the optimisation presented in Chapter 4. The participant feedback from our preference experiment in Chapter 5 brought to our attention that our colour-to-grey optimisation would give unnatural levels of contrast in some of our greyscales. Using a previously defined metric of variance/contract in the RGB, we could use our variance preserving solution on this 3-dimensional scenario. This would produce a single point (the position of which depends on the variance we wish in our greyscale) on a line segment that would give us the band weights for our contrast constrained greyscale. Further preference experiments would be needed to see if this satisfied users.

The natural extension to the work presented in Chapter 7 would be to investigate how much constraining the correlation (by prioritising the preservation of variance of the intensity mapping) had on the evaluation metrics. This would involve computing the band weights of maximum correlation (presented in Chapter 6) and then mapping them to the convex set (if the band weights define a point outside of it) that defines our mean preserving normalised image.

How our solution can be implemented on individual image segments is another avenue for further work. For example, using techniques based on the histogram of the image (i.e. Voronoi image segmentation) one could individually pan-sharpen vegetation, urban and oceanic scenes. Additional further work would be aimed at fine-tuning the image pre-preparation stages and lastly making comparisons with more recent and competitive hybrid pan-sharpening methods.

Bibliography

(2019). QuickBird spectral response. `https://dg-cms-uploads-production.s3.amazonaws.com/uploads/document/file/105/DigitalGlobe_Spectral_Response_1.pdf`.

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274 – 2281.

Aiazzi, B., Alparone, L., Baronti, S., and Garzelli, A. (2002). Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 40(10):2300 – 2312.

Aiazzi, B., Baronti, S., Lotti, F., and Selva, M. (2009). A comparison between global and context-adaptive pansharpening of multispectral images. *IEEE GeoScience and Remote Sensing Letters*, 6(2):302 – 306.

Aiazzi, B., Baronti, S., and Selva, M. (2007). Improving component substitution pansharpening through multivariate regression of MS+pan data. *IEEE Transaction on Geoscience and Remote Sensing*, 45(10):3230 – 3239.

Alparone, L., Facheris, L., Baronti, S., Garzelli, A., and Nencini, F. (2004). Fusion of multispectral and sar images by intensity modulation. In *Proceedings of the 7th International Conference on Image Fusion*, pages 637 – 643.

Alsam, A. (2009). Contrast enhancing colour to grey. In *Scandinavian Conference on Image Analysis*, volume 5575, pages 588 – 596.

Amro, I., Mateos, J., Vega, M., Molina, R., and Katsaggelos, A. K. (2011). A survey of classical methods and new trends in pansharpening of multispectral images. *Journal on Advances in Signal Processing*, 79:1 – 22.

Barsi, J. A., Lee, K., Kvaran, G., Markham, B. L., and Pedelty, J. A. (2014). The spectral response of the Landsat-8 operational land imager. *Remote Sensing*, 6:10232 – 10251.

Bhat, P., Curless, B., Cohen, M., and Zitnick, C. L. (2008). Fourier analysis of the 2D screened Poisson equation for gradient domain problems. In *European Conference for Computer Vision (ECCV)*, pages 114 – 128.

Boggione, G. A., Pires, E. G., Santos, P. A., and Fonseca, L. M. G. (2003). Simulation of a panchromatic band by spectral combination of multispectral ETM+ bands. In *Proceedings of International Symposium on Remote Sensing of Environmental*, pages 321 – 324.

Bolz, J., Farmer, I., Grinspun, E., and Schröder, P. (2003). Sparse matrix solvers on the GPU: Conjugate gradients and multigrid. *ACM Transactions on Graphics*, 22(3):917 – 924.

Boyd, S. and Vandenberghe, L. (2009). *Convex Optimisation.* Cambridge University Press.

Campbell, J. B. (2006). *Introduction to Remote Sensing.* Taylor & Francis, 4th edition.

Carper, W. J., Lillesand, T. M., and Kiefer, R. W. (1990). The use for intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing*, 56(4):459 – 467.

Chavez, P. S., Sides, S. C., and Anderson, J. A. (1991). Comparison of three different methods to merge multiresolution amd multispectral data: Landsat TM and SPOT panchromatic. *Photogrammetric Engineering & Remote Sensing*, 57(3):295 – 303.

Chen, S., Luo, J., Shen, Z., Luo, G., and Zhu, C. (2008). Using synthetic variable ratio method to fuse multi-source remoteley sensed images based on sensor spectral response. In *Geoscience and Remote Sensing Symposium*, pages 1084 – 1087. IEEE.

Choi, J., Yu, K., and Kim, Y. (2011). A new adaptive component-substitution-based satellite image fusion by using partial replacement. *IEEE Transaction on Geoscience and Remote Sensing*, 49(1):295 – 309.

Choi, M. (2006). A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter. *IEEE Transactions on GeoScience and Remote Sensing*, 44(6):1672 – 1682.

Choi, M., Kim, R. Y., Nam, M.-R., and Kim, H. O. (2005). Fusion of multispectral and panchromatic satellite images using the curvelet transform. *IEEE GeoScience and Remote Sensing Letters*, 2(2):136 – 140.

Connah, D., Finlayson, G. D., and Bloj, M. (2007). Seeing beyond luminance: a psychophysical comparison of techniques for converting colour images to greyscale. In *Proceedings of 15th Color Imaging Conference*, pages 336–341. IS&T and SID.

Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. Wiley.

C.R. Wylie, J. (2008). *Introduction to Projective Geometry*. Dover.

Cui, M., Hu, J., and Razdan, A. (2010). Color-to-gray conversion using ISOMAP. *The Visual Computer*, 26:1349–1360.

da Cunha, A. L., Zhou, J., and Do, M. N. (2006). The nonsubsampled contourlet transform: The nonsubsampled contourlet transform: Theory, design, and applications. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 15(10).

Davis, T. A. (2003). Umfpack version 4.1 user guide. Technical Report TR-03-008, University of Florida.

Dou, W. and Chen, Y. (2008). An improved IHS image fusion method with high spectral fidelity. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37:1253 – 1256.

Du, H., He, S., Sheng, B., Ma, L., and Lau, R. W. H. (2015). Saliency-guided color-to-gray conversion using region-based optimization. *IEEE Transactions on Image Processing*, 24(1):434 – 443.

Engeldrum, P. G. (2000). *Psychometric Scaling*, page 94. Imcotek Press.

Fattal, R., Lischinski, D., and Werman, M. (2002). Gradient domain high dynamic range compression. *ACM Transactions on Graphics*, 21(3):249 – 256.

Finlayson, G. D. and Matheson, T. N. (2012). A closed form solution for the brightness preserving colour to greyscale image conversion. In *Colour Imaging Conference*.

Fonseca, L., Namikawa, L., Castejon, E., Carvalho, L., Pinho, C., and Pagamisse, A. (2011). *Image Fusion for Remote Sensing Applications, Image Fusion and Its Applications.* InTech.

Fraleigh, J. B. and Beauregard, R. A. (1995). *Linear Algebra.* Addison-Wesley, 3rd edition.

Franco P. Preparata, M. I. S. (1985). *Computational Geometry an Introduction.* Springer-Verlag.

Garzelli, A., Nencini, F., and Capobianco, L. (2008). Optimal MMSE pan sharpening of very high resolution multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 46(1):228 – 236.

Gescheider, G. A. (1997). *Psychophysics the Fundamentals*, page 201. Lawrence Erlbaum Associates, 3rd edition.

Gharbia, R., Baz, A. H. E., Hassanien, A. E., and Tolba, M. F. (2014). Remote sensing image fusion approach based on Brovey and wavelets transforms. In *Proceedings of the Fith International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA)*, volume 303, pages 311 – 321.

Ghassemian, H. (2016). A review of remote sensing image fusion methods. *Information Fusion*, 32:75 – 89.

Gillespie, A. R., Kahle, A. B., and Walker, R. E. (1987). Color enhancement of highly correlated images. II. channel ratio and“chromaticity” transformation techniques. *Remote Sensing of Environment*, 22:343 – 365.

Golub, G. H. and Loan, C. F. V. (1996). *Matrix Computations*, page 149.

Gonzalez, R. C. and Woods, R. E. (2008). *Digital Image Processing.* Pearson Prentice Hall.

González-Audícana, M., Otazu, X., Fors, O., and Seco, A. (2005). Comparison between Mallat's and the 'à trous' discrete wavelet transform based algorithms for the fusion of multuspectral and panchromatic images. *International Journal of Remote Sensing*, 26(3):595 – 614.

González-Audícana, M., Saleta, J. L., Catalán, R. G., and García, R. (2004). Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition. *IEEE Transaction on Geoscience and Remote Sensing*, 42(6):1291 – 1299.

Gooch, A. A., Olsen, S. C., Tumblin, J., and Gooch, B. (2005). Color2gray: Salience-preserving color removal. *ACM Transactions on Graphics – Proceedings of ACM SIGGRAPH*, 24(3):634–639.

Grundland, M. and Dodgson, N. A. (2005). The decolorize algorithm for contrast enhancing, color to grayscale conversion. Technical Report 649, University of Cambridge Computer Laboratory.

Guo, L. J. and Moore, J. M. (1998). Pixel block intensity modulation: adding spatial detail to TM band 6 thermal imagery. *International Journal of Remote Sensing*, 19(13):2477 – 2491.

Guo, L. J., Moore, J. M., and Haigh, J. D. (1997). Simulated reflectance technique for ATM image enhancement. *International Journal of Remote Sensing*, 18(2):243 – 254.

Guo, Q., Chen, S., Leung, H., and Liu, S. (2010). Covariance intersection based image fusion technique with application to pansharpening in remote sensing. *Information Sciences*, 180:3434 – 3443.

Harris, M. D. C. (2015). *Crowd-Sourced Data and its Applications for New Algorithms in Photographic Imaging*. PhD thesis, CMP, University of East Anglia.

He, K., Sun, J., and Tang, X. (2013). Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6):1397 – 1409.

Holschneider, M. and Tchamitchian, P. (1989). *Regularité local de la function 'non-differentiable' the Riemann*. Paris: Springer-Verlag.

Hong, G. and Zhang, Y. (2003). High resolution image fusion based on wavelet and IHS transformations. In *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*.

Jameel, A., Riaz, M. M., and Ghafoor, A. (2016). Guided filter and IHS-based pan-sharpening. *IEEE Sensors Journal*, 16(1):192 – 194.

Jin, B. and Süsstrunk, S. (2017). Preserving perceptual contrast on decolorization with optimized color orders. In *IEEE International Conference on Image Processing (ICIP)*, pages 17 – 20.

Jordan, D. W. and Smith, P. (1997). *Mathematical Techniques: An introduction for the engineering, physical and mathematical sciences.* Oxford University Press, 2nd edition.

Kaplan, N. H., Erer, I., and Elibol, F. (2012). Fusion of multispectral and panchromatic images by combining bilateral filter and IHS transform. *20th European Signal Processing Conference (EUSIPCO 2012)*, (20):2501 – 2505.

Khan, M. M., Chanussot, J., Condat, L., and Montanvert, A. (2008). Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique. *IEEE GeoScience and Remote Sensing Letters*, 5(1):98 – 102.

Kim, Y., Jang, C., Demouth, J., and Lee, S. (2009). Robust color-to-gray via nonlinear global mapping. *ACM Transactions on Graphics (SIGGRAPH)*, 28(5).

Kuhn, G. R., Oliveira, M. M., and Fernandes, L. A. F. (2008). An improved contrast enhancing approach for color-to-grayscale mappings. *The Visual Computer*, 24(7-9):505–514.

Laben, C. A. and Brower, B. V. (2000). Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. *Patent 6,011,875.*

Leung, Y., Liu, J., and Zhang, J. (2014). An improved adaptive intensity-hue-saturation method for the fusion of remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 11(5):985 – 989.

Li, S., Kang, X., Fang, L., Hu, J., and Yin, H. (2017). Pixel-level image fusion: A survey of the state of the art. *Information Fusion*, 33:100 – 112.

Li, X., Xu, Q., and Gao, F. (2015). Pansharpening based on an improved ratio enhancement. *IEEE International Geoscience and Remote Sensing Symposium*, pages 1100 – 1103.

Ling, Y., Ehlers, M., Usery, E. L., and Madden, M. (2007). FFT-enhanced IHS transform method for fusing high-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61:381 – 392.

Liu, J. G. (2000). Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details. *International Journal of Remote Sensing*, 21(18):3461 – 3472.

Lu, C., Xu, L., and Jia, J. (2012a). Contrast preserving decolorization. *IEEE International Conference on Computational Photography*, pages 1 – 7.

Lu, C., Xu, L., and Jia, J. (2012b). Real-time contrast preserving decolorization. *SIGGRAPH Asia 2012 Technical Briefs*, (34).

Ma, J. and Plonka, G. (2010). The curvlet transform. In *IEEE Signal Processing Magazine*, volume 27. IEEE.

Mahyari, A. G. and Yazdi, M. (2011). Panchromatic and multispectral image fusion based on maximization of both spectral and spatial similarities. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):1976 – 1985.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(7):674 – 693.

Mascarenhas, N. D. A., Banon, G. J. F., and Fonseca, L. M. G. (1991). Simulation of a panchromatic band by spectral linear combination of multispectral bands. In *Proceedings of International Geoscience and Remote Sensing Symposium*, pages 321 – 324.

Mitianoudis, N., Tzimiropoulos, G., and Stathaki, T. (2010). Fast wavelet-based pansharpening of multi-spectral images. In *IEEE International Conference on Imaging Systems and Techniques (IST)*.

Montagna, R. (2011). *Path based computations in colour image processing.* PhD thesis, Computing Sciences, University of East Anglia.

Munechika, C. K., Warnick, J. S., Salvaggio, C., and Schott, J. R. (1993). Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering & Remote Sensing*, 59(1):67 – 72.

Nayatani, Y. (1997). Simple estimation methods for the Helmholtz-Kohrausch effect. *Color Research and Applications*, 22(6):385 – 401.

Neumann, L., Čadík, M., and Nemcsics, A. (2007). An efficient perception-based adaptive color to gray transformation. In *Proceedings of Computational Aesthetics 2007*, pages 73– 80. Eurographics Association.

Nowak, R. and Baraniuk, R. (1998). Adaptive weighted highpass filters using multiscale analysis. *IEEE Transactions on Image Processing*, 7(7):1068 – 1074.

Núñez, J., Otazu, X., Fors, O., Prades, A., c Palá, V., and Arbiol, R. (1999). Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1204 – 1211.

O'Rourke, J. (1998). *Computational Geometry in C.* Cambridge University Press, 2nd edition.

Pan, V. Y. and Chen, Z. Q. (1999). The complexity of the matrix eigenproblem. In *Proceedings of the thirty-first annual ACM symposium on Theory of Computing*, pages 507 – 516.

Pat S. Chavez, J. and Kwarteng, A. Y. (1989). Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogrammetric Engineering & Remote Sensing*, 55(3):339 – 348.

Pohl, C. and van Genderen, J. (2014). Remote sensing image fusion: an update in the context of digital earth. *International Journal of Diigital Earth*, 7(2):158 – 172.

Pohl, C. and van Genderen, J. L. (1998). Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, 19(5):823 – 854.

Price, J. C. (1987). Combining panchromatic and multispectral imagery from dual resolution satellite instruments. *Remote Sensing of Environment*, 21:119 – 128.

Qiu, M., Finlayson, G. D., and Qiu, G. (2008). Contrast maximizing and brightness preserving color to greyscale image conversion. In *Computer Graphics, Imaging and Visualization*, pages 347–351.

Rahmani, S., Strait, M., Merkurjev, D., Moeller, M., and Wittman, T. (2010). An adaptive IHS pan-sharpening method. *IEEE GeoScience and Remote Sensing Letters*, 7(4).

Ranchin, T. and Wald, L. (2000). Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, 66(1):49 – 61.

Ranchin, T., Wald, L., and Mangolini, M. (1996). The ARSIS method: a general solution for improving spatial resolution of images by means of sensor fusion. In *Proceedings of the Conference on Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*, pages 53 – 58.

Rasche, K., Geist, R., and Westall, J. (2005). Re-coloring images for gamuts of lower dimension. *Computer Graphics Forum*, 24(3):423–432.

Schott, J. R. (2007). *Remote Sensing The Image Chain Approach.* Oxford University Press, 2nd edition.

Schowengerdt, R. A. (1980). Reconstruction of multispatial, multispectral image data using spatial frequency content. *Photogrammetric Engineering & Remote Sensing*, 46(10):1325 – 1334.

Schowengerdt, R. A. (2007). *Remote Sensing: Models and Methods for Image Processing.* Elsevier, 3rd edition.

Shah, V. P., Younan, N. H., and King, R. L. (2008). An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5):1323 – 1335.

Shahdoosti, H. R. and Ghassemian, H. (2011). Multispectral and panchromatic image fusion by combining spectral PCA and spatial PCA methods. *Modares Journal of Electrical Engineering*, 11(3):19 – 27.

Shahdoosti, H. R. and Ghassemian, H. (2015). Fusion of MS and PAN images preserving spectral quality. *IEEE Geoscience and Remote Sensing Letters*, 12(3):611 – 615.

Shahdoosti, H. R. and Ghassemian, H. (2016). Combining the spectral PCA and spatial PCA fusion methods. *Information Fusion*, 27:150 – 160.

Shettigara, V. K. (1992). A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set. *Photogrammetric Engineering & Remote Sensing*, 58(5):561 – 567.

Smith, A. R. (1978). Color gamuts transform pairs. *Computer Graphics*, 12:12 – 19.

Smith, K., edouard Landes, P., Thollot, J., and Myszkowski, K. (2008). Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. In *Computer Graphics Forum*, volume 27, pages 193–200.

Song, M., Tao, D., Chen, C., Li, X., and Chen, C. W. (2010). Color to gray: Visual cue preservation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1537 – 1552.

Song, Y., Bao, L., Xu, X., and Yang, Q. (2013). Decolorization: Is rgb2gray() out? *Proceeing SIGGRAPH Asia Tecnhical Briefs*, (15).

Song, Y., Bao, L., and Yang, Q. (2014). Real-time video decolorization using bilateral filtering. *IEEE Conference on Applications of Computer Vision (WACV)*, pages 159 – 166.

Song, Y., Wu, W., Liu, Z., Yang, X., Liu, K., and Lu, W. (2016). An adaptive pansharpening method by using weighted least squares filter. *IEEE Geoscience and Remote Sensing Letters*, 13(1):18 – 22.

Stathaki, T., editor (2008). *Image Fusion: Algorithms and Aplications.* Elsevier, 1st edition.

Stephens, R. (2017). Calculate where a line and an ellipse intersect in C#. `http://csharphelper.com/blog/2017/08/calculate-where-a-line-segment-and-an-ellipse-intersect-in-c/`.

Sujitha, S. M. S., Selvathi, D., and Azees, M. (2013). Hybrid fusion approach of panchromatic and multispectral images using IHS and Radon transform. In *Proceedings of Information and Communication Technologies (ICT)*, pages 1309 – 1313. IEEE.

Thomas, C., Ranchin, T., Wald, L., and Chanussot, J. (2008). Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transaction on Geoscience and Remote Sensing*, 46(5):1301 – 1312.

Tian, N., Guo, L., Ren, M., and Ai, C. (2014). Implementingthe matrix inversion by Gauss-Jordan method with CUDA. In *International Conference on Wireless Algorithms, Systems, and Applications*, pages 44 – 53. Springer.

Tu, T.-M., Huang, P. S., Hung, C.-L., and Chang, C.-P. (2004). A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery. *IEEE Geoscience and Remote Sensing Letters*, 1(4):309–312.

Tu, T.-M., Su, S.-C., Shyu, H.-C., and Huang, P. S. (2001). A new look at IHS-like image fusion methods. *Information Fusion*, 2(3):177 186.

Valizadeh, S. A. and Ghassemian, H. (2012). Remote sensing image fusion using combining IHS and curvelet transform. *6th International Symposium on Telecommunications (IST'2012).*

Verlet, L. (1967). Computer "experiments" on classical fluids. i. thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159(1):98 – 103.

Vijayaraj, V., O'Hara, C. G., and Younan, N. H. (2004). Quality analysis of pansharpened images. In *Geoscience and Remote Sensing Symposium*, pages 85 – 88.

Vivone, G., Alparone, L., Chanussot, J., Mura, M. D., Garzelli, A., Licciardi, G. A., Restaino, R., and Wald, L. (2015). A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and RemoteSensing*, 53(5):2565 – 2586.

Wald, L., Ranchin, T., and Mangolini, M. (1997). Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric engineering and remote sensing*, 63(6):691 – 699.

Wang, L., Cao, X., and Chen, J. (2008). ISVR: an improved synthetic variable ratio method for image fusion. *Geocarto International*, 23(2):155 – 165.

Weisstein, E. W. (2019a). Circle-line intersection. `http://mathworld.wolfram.com/Circle-LineIntersection.html`.

Weisstein, E. W. (2019b). Ellipse-line intersections. `http://mathworld.wolfram.com/Ellipse-LineIntersection.html`.

W.H.Press, S.D.Teukolsk, W.T.Verterling, and B.P.Flannery (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2nd edition.

Wiemker, R., Prinz, B., Meister, G., Franck, R., and Spitzer, H. (1998). Accuracy assessment of vegetation monitoring with high spatial resolution satellite imagery - local, regional, global. In *International Symposium on Resource and Environmental Monitoring*.

Xu, J., Guan, Z., and Liu, J. (2008). An imporved IHS fusion method for merging multi-spectral and panchromatic images considering sensor spectral response. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B7):1169 – 1174.

Yang, Q., Tan, K.-H., and Ahuja, N. (2009). Real-time $O(1)$ bilateral filtering. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 557 – 564.

Yang, Y., Wan, W., Huang, S., Yuan, F., Yang, S., and Que, Y. (2016). Remote sensing image fusion based on adaptive IHS and multiscale guided filter. *IEEE Access*, 4.

YE, Y. and TSE, E. (1989). An extension of Karmarkar's projective algorithm for convex quadratic programming. *Mathematical Programming*, 44:157 – 179.

Yésou, H., Besnus, Y., and Rolet, J. (1993). Extraction of spectral information from Landsat TM data and merger with SPOT panchromatic imagery - a contribution to the study of geological structures. *Journal of Photogrammetry and Remote Sensing*, 48(5):23 – 36.

Yocky, D. A. (1996). Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data. *Photogrammetric Engineering & Remote Sensing*, 62(9):1067 – 1074.

Zhang, J. (2010). Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*, 1(1):5 – 24.

Zhang, Y. (1999). A new merging method and its spectral and spatial effects. *International Journal of Remote Sensing*, 20(10):2003 – 2014.

Zhang, Y. (2004). Understanding image fusion. *Photogrammetric Engineering & Remote Sensing*, pages 657 – 661.

Zhang, Y. and Hong, G. (2005). An IHS and wavelet integrated approach to improve pan-sharpening visual quality of natural colour IKONOS and QuickBird images. *Information Fusion*, 6:225 – 234.

Zhou, J., Civco, D. L., and Silander, J. A. (1998). A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International Journal of Remote Sensing*, 19(4):743 – 757.

Zhou, Z.-M., Yang, P.-L., Li, Y.-X., Yin, W., and Jiang, L. (2013). Joint IHS and variational methods for pan-sharpening of very high resolution imagery. *IEEE International Geoscience and Remote Sensing Symposium*, pages 2597 – 2600.

Zitová, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21:977 – 1000.